

Distribution rules of crystallographic systematic absences on the Conway topograph and their application to powder auto-indexing*

Ryoko Oishi-Tomiyasu[†]

January 4, 2013

Abstract

Powder auto-indexing is the crystallographic problem of lattice determination from an average theta series. There, in addition to all the multiplicities, the lengths of part of lattice vectors cannot be obtained owing to systematic absences. As a consequence, solutions are not always unique. We develop a new algorithm to enumerate powder auto-indexing solutions. This is a novel application of the reduction theory of positive-definite quadratic forms to a problem of crystallography. Our algorithm is proved to be effective for all types of systematic absences, using their newly obtained common properties. The properties are stated as distribution rules for lattice vectors corresponding to systematic absences on a topograph. Conway defined topographs for 2-dimensional lattices as graphs whose edges are associated with $|l_1|^2$, $|l_2|^2$, $|l_1 + l_2|^2$, $|l_1 - l_2|^2$, where l_1, l_2 are lattice vectors. In our enumeration algorithm, topographs are utilized as a network of lattice vector lengths. As a crystal structure is a lattice of rank 3, the definition of topographs is generalized to any higher dimensional lattices using Voronoi's second reduction theory. The use of topographs allows us to speed up the algorithm. The computation time is reduced to 1/250–1/32, when it is applied to real powder diffraction patterns. Another advantage of our algorithm is its robustness to missing or false elements in the set of lengths extracted from a powder diffraction pattern. Conograph is the powder indexing software which implements the algorithm. We present results of Conograph for 30 diffraction patterns, including some very difficult cases.

1 Introduction

Lattice determination problems have been the target of much interest in mathematics. In particular, many mathematicians worked on lattice determination from a theta series in the second half of the 20th century, with the aim of providing higher dimensional counterexamples for the isospectral problem proposed by Kac [16]. Lattice determination from a theta series was finally resolved in [25], [26]. However, it is not well known in the mathematics community that a similar problem was being studied in crystallography during this period.

The lattice determination problem in crystallography is called powder auto-indexing. There a set of lattice parameters is determined from a powder diffraction pattern. Powder auto-indexing is equivalent to lattice determination from an average theta series, as shown in Appendix A. At present, powder diffraction is the principal technique for

*This research was partly supported by a Grant-in-Aid for Young Scientists (B) (No. 22740077) and the Ibaraki prefecture (J-PARC-23D06).

[†]High energy accelerator research organization, Tsukuba, Japan. (ryoko.tomiyasu@kek.jp)

determining the structure of materials that do not necessarily form large-sized crystals. A highly automated system of analyzing powder diffraction patterns is required, both for basic scientific research and a wide range of industrial applications, not least in the pharmaceutical industry. Hence, it is necessary to establish a powerful and reliable powder auto-indexing algorithm and software.

In this paper, the powder auto-indexing problem is formulated and solved. This is regarded as a new application of the reduction theory of positive-definite quadratic forms to crystallography. In consideration of its application to crystallography, we mainly discuss cases when the lattice has a rank of $N = 2, 3$. Similar to a theta series, information about the lengths of lattice vectors is extracted from an average theta series (Section 2). The most notable difference is that it is very difficult to acquire the multiplicities of the lengths (*i.e.*, number of lattice vectors satisfying $|l|^2 = q$ for fixed $q \in \mathbb{R}_{>0}$) from an average theta series, and only a part of the lengths can be obtained owing to the crystallographic phenomenon of systematic absences.

Systematic absences have not been defined sufficiently explicitly as a mathematical notion. We provide a definition in Section 4. As in the International Tables [14], systematic absences are classified by the triplet (G, H, U) :

- isomorphism class of crystallographic group $G \subset O(N) \ltimes \mathbb{R}^N$,
- subgroup $H \subset G$, which contains the lattice L of G , *i.e.*, an abelian subgroup $L \simeq \mathbb{Z}^N$ consisting of all the translations in G ,
- a connected component U of $(\mathbb{R}^N/L)^H$.

Consequently, for fixed N , systematic absences are categorized by finitely many triplets (Proposition 4.1). When a periodic function \wp belongs to the type (G, H, U) , the lengths of any $l^* \in \Gamma_{ext}(G, H, U)$ is not extracted from the average theta series of \wp easily, where $\Gamma_{ext}(G, H, U)$ is a subset of the reciprocal lattice L^* of L determined from (G, H, U) .

To date, powder auto-indexing algorithms have been studied and improved by experts in crystallography. Among existing software packages, Ito [10], [31], TREOR (trial-and-error method [17]), and DICVOL (dichotomy method [4]) are widely used. McMaille (a grid search based on the Monte Carlo method [19]) and X-cell (dichotomy method [21]) were developed comparatively recently. With regard to existing powder auto-indexing software, several considerable problems have been reported. These include:

- existence of more than one solutions,
- systematic absences,
- missing or false elements in the set of extracted lengths,
- observation errors contained in the length values,
- large zero-point shift,
- overlapping peaks.

In some software, these problems are caused by limitations intended to suppress the computation time. Therefore, both computation time and success rate should be considered to discuss powder auto-indexing algorithms. (For example, although the simplest brute force grid search may obtain the highest success rate, it takes from a few hours to a few days to execute in monoclinic and triclinic cases.) This paper presents a method that can resolve each of the problems listed above. The first three are explained in more detail in the following discussion. The next two are resolved using the estimated error

$\text{Err}[|l|^2]$ of $|l|^2$ (see (A2) in Section 3 and the beginning of Section 7). The final problem is also resolved naturally in our algorithm (Section 8.1).

Our three main results are as follows. First, we formulate powder auto-indexing as a mathematical problem, and summarize mathematical results on the cardinality of solutions in powder auto-indexing. This is considered the most important foundation for the following discussion; in general, a lattice L is not determined uniquely in $N \geq 2$, even if the rank N of L and all the lengths $l \in L$ are given (Appendix B). On the other hand, the cardinality of L_2 satisfying $\Lambda_L = \Lambda_{L_2}$ is always finite in $N \leq 4$, therefore it is possible to enumerate all such L_2 algorithmically (Appendix C). However, it is not certain whether the number of solutions is actually finite in powder auto-indexing, owing to systematic absences and the finite observational range. Fortunately, using both physical constraints and mathematical theorems, we obtain a necessary condition for assuming the finiteness of solutions, except for events with zero probability (Section 3). As a result, it is only necessary to enumerate finitely many solutions for powder auto-indexing.

Our second result is a new algorithm to enumerate powder auto-indexing solutions (Sections 7.1 and 7.2). We prove that it works regardless of the type of systematic absences. Our basic idea is to use the graph whose edges are associated with a set of squares of lengths $|l_1^*|^2, |l_2^*|^2, |l_1^* + l_2^*|^2, |l_1^* - l_2^*|^2$ for some lattice vectors $l_1^*, l_2^* \in L^*$.

In Ito's method, the parallelogram law $2(|l_1^*|^2 + |l_2^*|^2) = |l_1^* + l_2^*|^2 + |l_1^* - l_2^*|^2$ is used to obtain Gram matrices of sublattices of rank 2 (called *zones*). However, simple use of the parallelogram law does not always provide successful results, as introduced in Fact 1 of Section 6.2. Therefore, instead of just enumerating a set of lengths satisfying the parallelogram law, our algorithm utilizes graphs whose edges are associated with $|l_1^*|^2, |l_2^*|^2, |l_1^* + l_2^*|^2$, and $|l_1^* - l_2^*|^2$ as a network of lattice vector lengths. This provides a comprehensive method of analyzing how the two sets $\{|l_1^*|^2, |l_2^*|^2, |l_1^* + l_2^*|^2, |l_1^* - l_2^*|^2\}$ and $\{|k_1^*|^2, |k_2^*|^2, |k_1^* + k_2^*|^2, |k_1^* - k_2^*|^2\}$ are related to each other.

For lattices of rank 2, such a graph has already been defined by Conway, who called it a topograph [6]. As explained in section 5.1, a graph having the required property is constructed for general N by using Voronoi's second reduction theory. We also call this a topograph. Basic properties of topographs for lattices of rank $N = 2, 3$ are explained in Section 5.2.

In Section 6, we shall see how primitive vectors of $l^* \in L^*$ belonging to $\Gamma_{\text{ext}}(G, H, U)$ are distributed on a topograph. Although the following theorem for the case of $N = 2$ has not been described explicitly, it provides a theoretical reason for the parallelogram law working appropriately for $N = 2$.

Theorem 1. *Let (G, H, U) be a type of systematic absence in $N = 2$, and let L^* be the reciprocal lattice of L , a lattice consisting of all the translations in G . Then, for any primitive vector l^* of L^* belonging to $\Gamma_{\text{ext}}(G, H, U)$, there exists $l_2^* \in L^*$ such that $l_1^* \cdot l_2^* = 0$ and l_1^*, l_2^* make a basis of L^* .*

In brief, Theorem 1 claims that $l^* \in \Gamma_{\text{ext}}(G, H, U)$ only appear in the gray area in Figure 1, regardless of the type (G, H, U) . With regard to the case $H = L$, we shall introduce a theoretical proof using topographs in Section 6.1. The case $H \supsetneq L$ is not proved here since it is verified easily by checking lists in [14].

By Theorem 1, a subgraph of a topograph with infinitely many edges is formed by unifying substructures as in Figure 2 associated with the lengths $|l_1^*|^2, |l_2^*|^2, |l_1^* + l_2^*|^2, |l_1^* - l_2^*|^2$ of $l_1^*, l_2^* \in L^* \setminus \Gamma_{\text{ext}}(G, H, U)$ that are easily extracted from an average theta series.

For $N = 3$, similar statements are proved in Theorems 3 and 4. In this case, it is necessary to use the following formula instead of the parallelogram law:

$$3|l_1^*|^2 + |l_1^* + 2l_2^*|^2 = 3|l_2^*|^2 + |2l_1^* + l_2^*|^2. \quad (1)$$

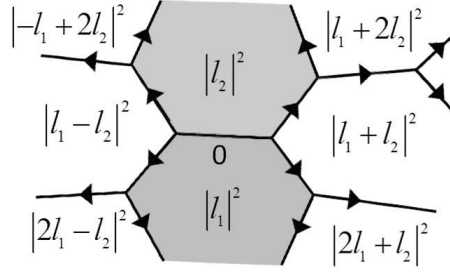


Figure 1: Reciprocal lattice vectors that are allowed to correspond to systematic absences.

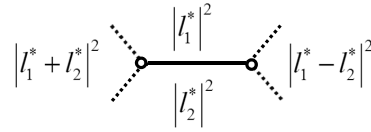


Figure 2: Substructure of a topograph corresponding to the parallelogram law.

As another consequence of Theorems 3 and 4, our algorithm enumerates the Gram matrices $(l_i^* \cdot l_j^*)_{1 \leq i, j \leq 3}$ for multiple bases l_1^*, l_2^*, l_3^* of the true solution L^* . This makes the enumeration procedure robust against missing or false elements in the set of extracted lengths, as explained in Section 7.4.

For the third result, we provide a method to speed up the enumeration process in Section 7.3. In order not to reduce the rate to acquire the true solution, Theorems 3 and 4 are used again here. In the test using actual powder diffraction patterns in Section 8.2, it is demonstrated that the improvement makes the enumeration speed 32–250 times faster. As a result, the enumeration process is executed in a few minutes at most.

The novel algorithm for lattices of rank $N = 3$ is implemented in the powder auto-indexing software Conograph. In section 8, we introduce the default parameters and results of Conograph. The default parameters are selected so that less-experienced users of the software can obtain good results without modifying them. We prepare 30 sets of test data, including difficult cases such as samples from Structure Determination by Powder Diffractometry Round Robin-2 (SDPDRR-2). Results for rather difficult cases are explained in Examples 3–6. The total time for powder auto-indexing did not exceed several minutes. The Conograph software is scheduled to be distributed in the near future from <http://sourceforge.jp/projects/conograph/>.

Notation and symbols

The notation and symbols used in this paper are summarized in this section. The inner product of the Euclidean space \mathbb{R}^N is denoted by $u \cdot v$, and the Euclidean norm $u \cdot u$ is denoted by $|u|^2$. The standard basis ${}^i(0, \dots, 0, \overset{i}{1}, 0, \dots, 0)$ of \mathbb{Z}^N is denoted by \mathbf{e}_i ($1 \leq i \leq N$).

A *lattice* L of rank N is a discrete and cocompact subgroup of \mathbb{R}^N . For any lattice L , there are linearly independent vectors $v_1, \dots, v_N \in \mathbb{R}^N$ over \mathbb{R} such that v_1, \dots, v_N generate L as a \mathbb{Z} -module. In this case, v_1, \dots, v_N are called a *basis* of L , and the matrix $(v_i \cdot v_j)_{1 \leq i, j \leq N}$ is called a *Gram matrix* of L . The *reciprocal lattice* L^* of L is defined as $L^* := \{l^* \in \mathbb{R}^N : l \cdot l^* \in \mathbb{Z} \text{ for all } l \in L\}$.

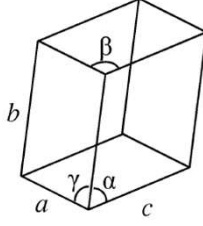


Figure 3: Lattice parameters and the unit cell.

If $\{v_1, \dots, v_i\} \subset L$ ($1 \leq i < N$) is extended to a basis of L , it is called a *primitive set* of L . In particular, $v \in L$ is a *primitive vector* of L if and only if $\{v\}$ is a primitive set of L . $P_n(L)$ is the set consisting of all the primitive sets of L of cardinality n .

A symmetric matrix $(s_{ij})_{1 \leq i, j \leq N}$ is always identified with a *quadratic form* $Q(x_1, \dots, x_N) := \sum_{i=1}^N \sum_{j=1}^N s_{ij} x_i x_j$. The linear space consisting of $N \times N$ symmetric matrices with real entries is denoted by \mathcal{S}^N . $\mathcal{S}_{>0}^N$ (resp. $\mathcal{S}_{\geq 0}^N$) is the subset of \mathcal{S}^N consisting of all the positive-definite (resp. semidefinite) symmetric matrices. $S_1, S_2 \in \mathcal{S}^N$ are *equivalent* if and only if there exists $g \in GL_N(\mathbb{Z})$ such that $S_2 = g S_1 g^t$.

For any $S \in \mathcal{S}^N$, elements of Λ_S are called *representations* of S over \mathbb{Z} .

$$\Lambda_S := \{ {}^t v S v : 0 \neq v \in \mathbb{Z}^N \}. \quad (2)$$

In crystallography, a crystal lattice in three-dimensional Euclidean space is represented by a set of lattice parameters a, b, c, α, β , and γ as in Figure 3. This parameterization is transformed into a 3×3 matrix $S := (s_{ij})_{1 \leq i, j \leq 3}$ as follows, and S is the Gram matrix of the *Bravais lattice* of the crystal.

$$\begin{aligned} s_{11} &= a^2, & s_{22} &= b^2, & s_{33} &= c^2, \\ s_{12} &= ab \cos \gamma, & s_{13} &= ac \cos \beta, & s_{23} &= bc \cos \alpha. \end{aligned} \quad (3)$$

In general, for any lattice $L \subset \mathbb{R}^N$, its automorphism group is defined by $\{g \in GL(3, \mathbb{Z}) : g S g^t = S\}$, and L is categorized into its Bravais type by the conjugacy class of the group in $GL(3, \mathbb{Z})$. In $N = 3$, it is known that there exist 14 Bravais types. In crystallography, selection of the Bravais lattice $L_2 \subset L$ and its basis l_1, l_2, l_3 is standardized according to the Bravais type of L . S in (3) is the Gram matrix defined for the basis of L_2 . When L belongs to the category called a primitive centring, $L = L_2$, and l_1, l_2, l_3 are chosen by the method called the Niggli reduction, which is very similar with the Minkowski reduction in $N = 3$ (see [14]). When L belongs to the category called a face-centered (resp. body-centered) lattice, there exist $l_1, l_2, l_3 \in L$ such that L is generated by $\frac{l_i + l_j}{2}$ ($1 \leq i < j \leq 3$) (resp. $\frac{l_1 + l_2 + l_3}{2} - l_i$ ($1 \leq i \leq 3$)), and $l_i \cdot l_j = 0$ ($1 \leq i < j \leq 3$) holds. Regardless of the choice of l_1, l_2, l_3 , their Gram matrix S and L_2 generated by these l_1, l_2, l_3 are determined uniquely. With regard to the Bravais lattice, the explanation above is sufficient in order to understand our following discussions.

In the context of powder structure analysis, representations of S^{-1} over \mathbb{Z} are called *q-values* of diffraction peaks.

2 Outline of the powder auto-indexing problem

A function \wp on \mathbb{R}^N is said to be *periodic* if $L := \{l \in \mathbb{R}^N : \wp(x + l) = \wp(x) \text{ for any } x \in \mathbb{R}^N\}$ is a lattice. We call L the *period lattice* of \wp . For example, an electron (or nucleus)

density in a crystal (*cf.* the left figure in Figure 4) is a periodic function of $N = 3$. The following \wp provides its standard model.

$$\wp(x) = \sum_{i=1}^m \sum_{k=1}^{d_i} \sum_{l \in L} p_i(x - x_{ik} - l), \quad (4)$$

where

- m : number of different elements in a crystal,
- d_i : number of the i th element in the primitive cell \mathbb{R}^3/L ,
- $p_i(x)$: rapidly decreasing function on \mathbb{R}^3 that represents the electron distribution of respective atoms.

According to diffraction theory, if a crystal has an electron density \wp , the diffraction image of its single-crystal sample equals $cf_{single}(x^*; \wp)$ for some $c > 0$, where $f_{single}(x^*; \wp)$ is a sum of delta functions given by

$$f_{single}(x^*; \wp) := \sum_{l^* \in L^*} |\hat{\wp}(l^*)|^2 \delta(x^* - l^*), \quad (5)$$

$$\hat{\wp}(l^*) := \int_{\mathbb{R}^N/L} \wp(x) e^{-2\pi\sqrt{-1}x \cdot l^*} dx. \quad (6)$$

A powder sample is an ensemble of a very large number of randomly oriented crystallites. As a result, its diffraction image is proportional to the integration of $f_{single}(x^*; \wp)$ on a sphere of radius \sqrt{q} :

$$f_{powder}(q; \wp) := \int_{|x^*|^2=q} f_{single}(x^*; \wp) dx^* = 2\sqrt{q} \sum_{l^* \in L^*} |\hat{\wp}(l^*)|^2 \delta(q - |l^*|^2). \quad (7)$$

The $\hat{\wp}(l^*)$ is called a *structure factor* in crystallography.

The right figure in Figure 4 presents an actual powder diffraction pattern. It is obtained by replacing every delta function in (7) with some kind of peak-shape model function g that is close to a Gaussian distribution and satisfies $\int_{\mathbb{R}} g(q) dq = 1$.

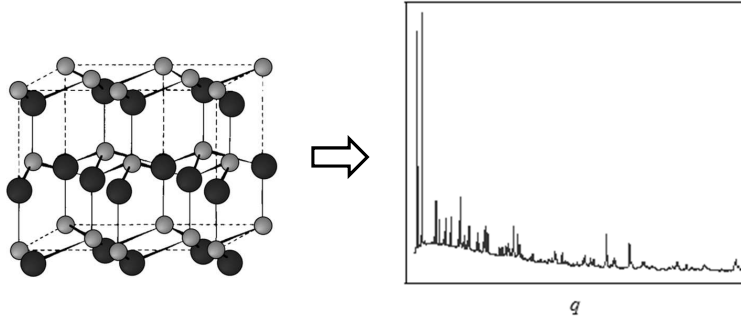


Figure 4: Powder diffraction pattern.

Ab-initio powder crystal structure determination retrieves the electron density \wp from the powder diffraction pattern under the assumption that the following additional information is available.

- chemical formula (*i.e.*, the ratio $[d_1 : \dots : d_m]$),
- density of a single crystal,

- rapidly decreasing functions p_i ($1 \leq i \leq m$).

Powder auto-indexing is the initial stage of ab-initio powder crystal structure determination, and aims to find the period lattice L of \wp . As the positions of the delta functions, elements of Λ_\wp are extracted from a powder diffraction pattern $f_{\text{powder}}(q; \wp)$.

$$\Lambda_\wp := \{|l^*|^2 : l^* \in L^*, F_\wp(q) \neq 0\}. \quad (8)$$

L is normally determined from elements of Λ_\wp in powder auto-indexing. After L is obtained, using the coefficients $F_\wp(q)$, powder crystal structure determination is carried out.

$$F_\wp(q) := \sum_{l^* \in L^*, |l^*|^2 = q} |\hat{\wp}(l^*)|^2. \quad (9)$$

3 Formulation of the powder auto-indexing problem

In this section, we formulate the problem explicitly. In powder auto-indexing, extracted Λ^{obs} is different from the set Λ^\wp of true $|l^*|^2$ ($l^* \in L^*$) owing to observational problems. Consequently, the Gram matrix S of the period lattice L of \wp must be retrieved from Λ^{obs} under the following assumptions.

- (A1) The observed range of a powder diffraction pattern is contained in a finite interval $[q_{min}, q_{max}] \subset (0, \infty)$. Consequently, only information about $\Lambda_\wp \cap [q_{min}, q_{max}]$ is available.
- (A2) Every $q^{obs} \in \Lambda^{obs}$ has some observation error. It may be assumed that the threshold $\text{Err}[q^{obs}]$ on the maximum error is given. (A method to compute $\text{Err}[q^{obs}]$ is introduced in [23].)
- (A3) Owing to mistakes and errors in acquisition of the peak-positions (*i.e.*, peak-search), Λ^{obs} differs from the true $\Lambda_\wp \cap [q_{min}, q_{max}]$. There exist small $\epsilon_1 > 0$ and $\epsilon_2 > 0$ satisfying the following conditions:
 - (i) For arbitrarily fixed $q \in \Lambda_\wp \cap [q_{min}, q_{max}]$, we have $q \notin \bigcup_{q^{obs} \in \Lambda^{obs}} [q^{obs} - \text{Err}[q^{obs}], q^{obs} + \text{Err}[q^{obs}]]$ with probability ϵ_1 .
 - (ii) For arbitrarily fixed $q^{obs} \in \Lambda^{obs}$, we have $\Lambda_\wp \cap [q^{obs} - \text{Err}[q^{obs}], q^{obs} + \text{Err}[q^{obs}]] = \emptyset$ with probability ϵ_2 .
- (A4) When the distribution \wp is constrained by group symmetry, infinitely many $\hat{\wp}(l^*)$ and $F_\wp(q)$ become zero owing to *systematic absences* (Section 4).

As assumed in (3), Λ_{obs} extracted from a powder diffraction pattern has missing or false elements. These are caused by observational problems including background noise or false peaks due to sample impurity.

With regard to (4), sometimes $F_\wp(q) = 0$ holds due to a special arrangement of atom positions x_{ik} , rather than systematic absences (*cf.* [12], [29]). However, the probability is zero if every x_{ik} is distributed uniformly in \mathbb{R}^N/L . (The only known exceptions are systematic absences.)

For the special arrangement with zero probability, we replace Λ_\wp by $\Lambda_{ext}(\wp)$ and consider (A3) instead of (A4):

$$\Lambda_{ext}(\wp) := \Lambda_\wp \cup \left\{ |l^*|^2 : \begin{array}{l} l^* \in L^*, F_\wp(q) = 0 \text{ owing to reasons} \\ \text{other than systematic absences} \end{array} \right\}. \quad (10)$$

($\tilde{A}3$) Assume that

- (i) For any $q \in \Lambda_{ext}(\wp) \cap [q_{min}, q_{max}]$, $q \notin \bigcup_{q^{obs} \in \Lambda^{obs}} [q^{obs} - \text{Err}[q^{obs}], q^{obs} + \text{Err}[q^{obs}]]$ occurs with probability ϵ_1 .
- (ii) For any $q^{obs} \in \Lambda^{obs}$, $\Lambda_{ext}(\wp) \cap [q^{obs} - \text{Err}[q^{obs}], q^{obs} + \text{Err}[q^{obs}]] = \emptyset$ occurs with probability ϵ_2 .

Under the conditions (A1)–(A4), infinitely many solutions may exist in some cases (for example, consider the case of very small $[q_{min}, q_{max}]$). Hence, additional assumptions are necessary in order to guarantee a finite number of solutions.

(A5) Owing to repulsive force between atoms, we may assume $\min\{|l|^2 : 0 \neq l \in L\} \geq d^2$ for a positive constant $d \approx 2\text{\AA}$. Then, by the inequalities on successive minima of L and its reciprocal lattice L^* proved by Lagarias *et al.* [18], in $N = 2, 3$, the maximum diagonal entry D_N of a Minkowski-reduced (defined in Appendix C) Gram matrix S of L^* satisfies

$$D_N \leq \frac{N+3}{4d^2} \max\{\gamma_i : 1 \leq i \leq N\}, \quad (11)$$

where γ_i is the Hermite constant:

$$\gamma_i := \sup \left\{ \min\{^t v S v : 0 \neq v \in \mathbb{Z}^N\} : S \in \mathcal{S}_{>0}^i, \det S = 1 \right\}. \quad (12)$$

In particular, $D_2 \leq \frac{5}{3}d^{-2}$ and $D_3 \leq 3 \cdot 2^{-1/3}d^{-2}$ follow from $\gamma_1 = 1$, $\gamma_2^2 = \frac{4}{3}$, $\gamma_3^3 = 2$. (In $N = 2$, the estimation is improved up to $D_2 \leq \frac{4}{3}d^{-2}$ easily.)

(A6) The length of the interval $\sqrt{q_{max}} - \sqrt{q_{min}}$ is sufficiently greater than $\sqrt{D_N}$ so that Λ^{obs} includes $|l_1^* \pm l_2^*|^2$, $|l_1^* \pm l_3^*|^2$, $|\pm l_1^* + l_2^* + l_3^*|^2$, and at least one of:

- (i) $|l_1^*|^2, |l_i^*|^2$,
- (ii) $|l_1^*|^2, |l_1^* + 2l_i^*|^2$,
- (iii) $|l_i^*|^2, |2l_1^* + l_i^*|^2$.

for both $i = 1, 2$, where l_1^*, l_2^*, l_3^* is some basis of L^* .

In (A5), 2\AA is selected as the minimum distance between the two closest lattice points of L satisfied by any existing crystals. (A6) looks rather artificial. This assumption is necessary for our algorithm in Table 5. By Theorems 3 and 4, (A6) holds except for events with zero probability if a sufficiently large q_{max} is chosen, regardless of the type of systematic absences. Under assumptions (A5) and (A6), the number of solutions is always finite, because $\Lambda^{obs} \subset [q_{min}, q_{max}]$ contains only finite elements and the Gram matrix of L^* is computed from a combination of elements of Λ^{obs} due to assumption (A6).

Here, it is still non-trivial how to select q_{max} . At least, it is clear $q_{max} \geq D_3$ is required; otherwise $\Lambda_{ext}(\wp) \cap [q_{min}, q_{max}]$ contain only $|l^*|^2$ of $l^* \in L_2$, some sublattice of L of rank less than 3 in some cases. On the other hand, as seen in (3) in Section 8, too many q -values are frequently extracted from the interval $[q_{min}, D_3]$ when we set $d = 2\text{\AA}$, hence $D_3 := 3 \cdot 2^{-1/3}d^{-2}$, nevertheless powder auto-indexing is very frequently successful even with a smaller interval. Since the time of our enumeration algorithm is roughly proportional to the fourth power of the number of elements of Λ^{obs} (*cf.* Section 7.3), the time can be significantly decreased by minimizing the range $[q_{min}, q_{max}]$. Considering the current accuracy of diffractometers and the power of personal computers, q_{max} should be chosen empirically to some degree, in addition to the theoretical estimation above. See (3) in Section 8.1 for a more detailed approach to this issue.

4 Summary of crystallographic groups and systematic absences

We now give some definitions for crystallographic groups and systematic absences. In the following, we represent the elements $x \in \mathbb{R}^N/L$ as a row vector and $l^* \in L^*$ as a column vector. Furthermore, any group action on \mathbb{R}^N/L (resp. L^*) is represented as a right action (resp. left action).

Any congruent transformation of the Euclidean space \mathbb{R}^N is represented as a composition of the orthogonal group $O(N)$ and a translation; if σ is a congruent transformation of \mathbb{R}^N , there exist $\tau \in O(N)$ and $\nu \in \mathbb{R}^N$ such that:

$$x^\sigma = x^\tau + \nu \text{ for any } x \in \mathbb{R}^N. \quad (13)$$

Such a σ is denoted by $\{\tau|\nu\}$. The group consisting of all congruent transformations of \mathbb{R}^N is the semidirect group $O(N) \ltimes \mathbb{R}^N$. By expanding the composition $(x^{\{\tau_1|\nu_1\}})^{\{\tau_2|\nu_2\}}$, it is seen that the group multiplication is given by:

$$\{\tau_1|\nu_1\} \cdot \{\tau_2|\nu_2\} = \{\tau_1\tau_2|\nu_1^{\tau_2} + \nu_2\}. \quad (14)$$

Definition 4.1. A crystallographic group is a discrete and cocompact subgroup of $O(N) \ltimes \mathbb{R}^N$.

A crystallographic group is also called a *wallpaper group* in $N = 2$, and a *space group* in $N = 3$.

For a crystallographic group G , two groups R_G and L are defined by:

$$R_G := \{\tau \in O(N) : \{\tau|\nu\} \text{ for some } \nu \in \mathbb{R}^N\}, \quad (15)$$

$$L := \{\nu \in \mathbb{R}^N : \{1_N|\nu\} \in G\}, \quad (16)$$

where 1_N is the identity of $O(N)$. R_G is called a *point group* of G . G is represented as the semidirect $G \simeq R_G \ltimes L$. From the definition of L , for any $\{\tau|\nu_\tau\} \in G$, the class $\nu_\tau + L \in \mathbb{R}^N/L$ is uniquely determined. Furthermore, the map $R_G \rightarrow \mathbb{R}^N/L : \tau \mapsto \nu_\tau + L$ is a 1-cocycle, *i.e.*, it satisfies

$$\nu_{\sigma\tau} \equiv \nu_\sigma^\tau + \nu_\tau \text{ mod } L. \quad (17)$$

The following was proved by Bieberbach.

Theorem 2 (Bieberbach [2], [3]). *In each N ,*

- (1) *L is a lattice and R_G is a finite subgroup of $O(N)$ consisting of τ that maps any elements of L to L .*
- (2) *There are only finitely many isomorphism classes of crystallographic groups.*
- (3) *A homomorphism $\varphi : G_1 \rightarrow G_2$ is an isomorphism between two crystallographic groups G_1 and G_2 if and only if there is an affine map α of \mathbb{R}^N such that $\varphi(g) = \alpha g \alpha^{-1}$.*

We now proceed to the definition of systematic absences. Let $G = R_G \ltimes L$ be a crystallographic group, and consider the following periodic function \wp :

$$\wp(x) = \sum_{i=1}^m \sum_{\sigma \in G} p_i(x - x_i^\sigma). \quad (18)$$

Note that the density model function \wp in (4), is represented as in (18) for some space group G .

Let $L^2(\mathbb{R}^N)$ be the L^2 -space, i.e., the set of all measurable functions f on \mathbb{R}^N with a finite L^2 -norm $\|f\|_2 := (\int_{\mathbb{R}^N} |f(x)|^2 dx)^{1/2} < \infty$. In order to compute the same list as [14], let us assume that

(Isotropy condition) p_i belongs to $L^2(\mathbb{R}^N)^{R_G}$, i.e., $p_i(x^\tau) = p_i(x)$ holds for any $\tau \in R_G$.

Then, $\wp(x^\sigma) = \wp(x)$ holds for any $\sigma \in G$, and \wp has the Fourier coefficient

$$\begin{aligned} \hat{\wp}(l^*) &= \sum_{i=1}^m \sum_{\sigma \in G} \int_{\mathbb{R}^N/L} p_i(x - x_i^\sigma) e^{-2\pi\sqrt{-1}x \cdot l^*} dx \\ &= \sum_{i=1}^m \hat{p}_i(l^*) \sum_{\sigma \in G/L} e^{-2\pi\sqrt{-1}x_i^\sigma \cdot l^*}, \end{aligned} \quad (19)$$

where $\hat{p}_i(l^*) := \int_{\mathbb{R}^N} p_i(x) e^{-2\pi\sqrt{-1}x \cdot l^*} dx$.

Hence, the probability of $\hat{\wp}(l^*) = 0$ depends on the size of $W_{G,l^*} \subset V := \mathbb{R}^N/L$.

$$W_{G,l^*} := \left\{ x \in V : \sum_{\sigma \in G/L} e^{-2\pi\sqrt{-1}x^\sigma \cdot l^*} = 0 \right\}. \quad (20)$$

Definition 4.2. For any subgroup of $L \subset H \subset G$, let V^H be the subset of V consisting of all the fixed points of H , and $U \subset V^H$ be a connected component of V^H . Under this notation, $\Gamma_{ext}(G)$ and $\Gamma_{ext}(G, H, U)$ are defined by

$$\Gamma_{ext}(G) := \{l^* \in L^* : V = W_{G,l^*}\}, \quad (21)$$

$$\Gamma_{ext}(G, H, U) := \{l^* \in L^* : U \subset W_{G,l^*}\}. \quad (22)$$

According to the terminology of crystallography, we say that $l^* \in L^*$ corresponds to systematic absences at general positions (resp. special positions) if and only if l^* belongs to $\Gamma_{ext}(G)$ (resp. $\Gamma_{ext}(G, H, U)$).

Remark 1. An interpretation of systematic absences is provided by the Laplacian operator on V ; the Laplacian operator $\Delta := \sum_{i=1}^N \frac{\partial^2}{\partial x_i^2}$ on $V := \mathbb{R}^N/L$ acts on $C^\infty(V)$, the set of smooth functions (i.e., derivatives of all orders are continuous) on V . Δ also acts on

$$C^\infty(V)^G := \{f \in C^\infty(V) : f(x) = f(x^\sigma) \text{ for any } x \in V \text{ and } \sigma \in G\}. \quad (23)$$

Recall that $\Lambda_{L^*} := \{|l^*|^2 : 0 \neq l^* \in L^*\}$ coincides with the set of non-zero eigenvalues of Δ for the action. (This is seen from the fact that any $f(x) \in C^\infty(V)$ is represented as a Fourier series $\sum_{l^* \in L^*} c_{l^*} e^{2\pi\sqrt{-1}x \cdot l^*}$, and we have $\Delta f(x) = \sum_{l^* \in L^*} |l^*|^2 c_{l^*} e^{2\pi\sqrt{-1}x \cdot l^*}$.) Similarly, $\Lambda_{ext}(G)$ is the set of non-zero eigenvalues for the action of Δ on $C^\infty(V)^G$, because $f \in C^\infty(V)^G$ is represented as $\sum_{[l^*] \in R_G \setminus L^*} c_{[l^*]} \sum_{\sigma \in G/L} e^{2\pi\sqrt{-1}x^\sigma \cdot l^*}$. $\Lambda_{ext}(G, H, U)$ is the subset of $\Lambda_{ext}(G)$ corresponding to an eigenvector $f \in C^\infty(V)^G$ that is not constantly zero on U .

In the following, we shall focus on $\Gamma_{ext}(G, H, U)$, because $\Gamma_{ext}(G) = \Gamma_{ext}(G, L, V)$. From the definition, it is clear that $-l^*$ and τl^* belong to $\Gamma_{ext}(G, H, U)$ for any $\tau \in R_G$ if and only if l^* does.

Let R_H be the image of H by the map $G \longrightarrow R_G : \{\tau | \nu_\tau\} \mapsto \tau$. If the image of V by $v \mapsto \sum_{\tau \in R_H} v^\tau$ is denoted by $V^{\sum_{R_H}}$, $V^{\sum_{R_H}}$ is open and closed in V^{R_H} . Hence the connected component including $x \in V^H$ is represented as $x + V^{\sum_{R_H}}$. Therefore, we may denote (22) by

$$\Gamma_{ext}(G, H, x + V^{\sum_{R_H}}) := \{l^* \in L^* : x + V^{\sum_{R_H}} \subset W_{G,l^*}\}. \quad (24)$$

When $H_i \subset G$ is the stabilizer subgroup of $x_i \in V$, it is not difficult to confirm that the following equivalence condition holds for \wp in (18) satisfying the isotropy condition.

$$\begin{aligned} \hat{\wp}(l^*) = 0 \text{ holds constantly when } (x_1, \dots, x_m, p_1, \dots, p_m) \\ \text{is perturbed in an open neighborhood in } V^{H_1} \times \dots \times V^{H_m} \times (L^2(\mathbb{R}^N)^{R_G})^m \\ \iff l^* \in \bigcap_{i=1}^m \Gamma_{ext}(G, H_i, x_i + V^{\Sigma_{R_{H_i}}}). \end{aligned} \quad (25)$$

The following Propositions and Corollaries summarize basic properties of systematic absences used in the subsequent sections. Proposition 4.1 states a well-known fact:

Proposition 4.1. *All types of systematic absences are classified by triplets $(G, H, x + V^{\Sigma_{R_H}})$, where G , H , and x range respectively in*

- (a) *isomorphism classes of a crystallographic group $G = R_G \ltimes L$,*
- (b) *subgroups $L \subset H \subset G$,*
- (c) *representatives of $V^{R_H}/V^{\Sigma_{R_H}}$.*

As a result, there are only finitely many types of systematic absences for each $N > 0$.

Proof. When $x \in V^H$ is fixed, the map $y \mapsto x + y + V^{\Sigma_{R_H}}$ gives a one-to-one correspondence between $V^{R_H}/V^{\Sigma_{R_H}}$ and connected components of V^H . Therefore, systematic absences are classified by the above three terms. The finite cardinality of (a) is obtained by the Bieberbach theorem, and (b) is straightforward because $[G : L] < \infty$ holds. In order to prove (c) is finite, we decompose V^{R_H} into a disjoint union $\bigcup_{y \in V^{R_H}/V^{\Sigma_{R_H}}} y + V^{\Sigma_{R_H}}$. Then V^{R_H} is closed in V , therefore compact. Since $y + V^{\Sigma_{R_H}}$ is open, the number of representatives of $V^{R_H}/V^{\Sigma_{R_H}}$ is required to be finite. \square

$\Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$ is computed using the following proposition.

Proposition 4.2. *For fixed $l^* \in L^*$, the equivalence relation among the right cosets $R_H \backslash R_G$ is defined by:*

$$R_H \tau_1 \overset{l^*}{\sim} R_H \tau_2 \iff \sum_{\tau \in R_H \tau_1} \tau l^* = \sum_{\tau \in R_H \tau_2} \tau l^*. \quad (26)$$

Then, for fixed $x \in V^H$, $l^* \in L^*$ belongs to $\Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$ if and only if the following holds:

$$\sum_{R_H \tau_2 \overset{l^*}{\sim} R_H \tau_1} e^{2\pi\sqrt{-1}x \cdot \{\tau_2 | \nu_{\tau_2}\} \cdot l^*} = 0 \text{ for every } R_H \tau_1 \in R_H \backslash R. \quad (27)$$

Proof. From (24), the following condition is equivalent to $l^* \in \Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$:

$$\sum_{R_H \tau_1 \in R_H \backslash R} e^{2\pi\sqrt{-1}((x+\delta x)^\tau + \nu_\tau) \cdot l^*} = 0 \text{ for any } \delta x \in V^{\Sigma_{R_H}}. \quad (28)$$

Furthermore,

$$\begin{aligned} \delta x^{\tau_1} \cdot l^* = \delta x^{\tau_2} \cdot l^* \text{ for any } \delta x \in V^{\Sigma_{R_H}} &\iff \sum_{\tau \in R_H} (\delta \tilde{x}^{\tau \tau_1} - \delta \tilde{x}^{\tau \tau_2}) \cdot l^* = 0 \text{ for any } \delta \tilde{x} \in V \\ &\iff \sum_{\tau \in R_H} \tau(\tau_1 l^* - \tau_2 l^*) = 0 \iff R_H \tau_1 \overset{l^*}{\sim} R_H \tau_2. \end{aligned} \quad (29)$$

Hence, (28) holds if and only if, for any $\delta x \in V^{\Sigma_{R_H}}$, we have:

$$\sum_{[\tau_1] \in (R_H \setminus R)/\sim} e^{2\pi\sqrt{-1}\delta x^{\tau_1} \cdot l^*} \sum_{R_H \tau_2 \sim_{l^*} R_H \tau_1} e^{2\pi\sqrt{-1}(x^{\tau_2} + \nu_{\tau_2}) \cdot l^*} = 0, \quad (30)$$

which leads to the statement. \square

Corollary 4.1. *Let M be the order of R_G , and $\mathcal{H}_{G,H} \subset \mathbb{R}^N$ be the following union of finite linear subspaces of dimension less than N :*

$$\mathcal{H}_{G,H} := \bigcup_{\substack{R_H \tau_1, R_H \tau_2 \in R_H \setminus R_G, \\ \sum_{\tau \in R_H} \tau(\tau_1 - \tau_2) \neq 0}} \{x^* \in \mathbb{R}^N : \sum_{\tau \in R_H} \tau(\tau_1 - \tau_2)x^* = 0\}. \quad (31)$$

There then exists $\Omega \subset L^/ML^*$ such that $l^* \in \Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}}) \iff l^* + ML^* \in \Omega$ holds for any $l^* \in L^* \setminus \mathcal{H}_{G,H}$.*

Proof. We may assume $x = 0$ from the translation $v \mapsto v - x$ in V . Then, $\nu_\tau = 0$ for any $\tau \in H$ follows from $x \in V^H$. Hence, $[\nu_\tau] \in H^1(G, V)$ are mapped to 0 by $H^1(G, V) \xrightarrow{\times M/m} H^1(G, V)$, where m is the order of R_H (cf. [28], Proposition 6 of Chap. VII). Thus, any $\tau \in R_G$ satisfies $\nu_\tau = y - y^\tau + \mu_\tau$ for some $y \in V$ and $\mu_\tau \in (M/m)^{-1}L/L$. Since $0 = \frac{M}{m}y - \frac{M}{m}y^\tau$ for any $\tau \in R_H$, we have $\frac{M}{m}y \in V^{R_H}$, and therefore $\sum_{\sigma \in R_H} \frac{M}{m}y^\sigma = My \in V^{\Sigma_{R_H}}$. Consequently, $y = \sum_{\sigma \in R_H} u^\sigma + \xi$ for some $u \in V$ and $\xi \in M^{-1}L/L$. Hence, $\nu_\tau = \sum_{\sigma \in R_H} u^\sigma - \sum_{\sigma \in R_H} u^{\sigma\tau} + \lambda_\tau$, where $\lambda_\tau = \mu_\tau + \xi - \xi^\tau \in M^{-1}L/L$. By Proposition 4.1, $l^* \in L^* \setminus \mathcal{H}_{G,H}$ belongs to $\Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$ if and only if the following holds for any $R_H \tau_1 \in R_H \setminus R_G$:

$$\begin{aligned} \sum_{\substack{R_H \tau_2 \in R_H \setminus R_G, \\ \sum_{\tau \in R_H} \tau(\tau_1 - \tau_2) = 0}} e^{2\pi\sqrt{-1}\nu_{\tau_2} \cdot l^*} &= \sum_{\substack{R_H \tau_2 \in R_H \setminus R_G, \\ \sum_{\tau \in R_H} \tau(\tau_1 - \tau_2) = 0}} e^{2\pi\sqrt{-1}(\sum_{\sigma \in R_H} u^\sigma - \sum_{\sigma \in R_H} u^{\sigma\tau_2} + \lambda_{\tau_2}) \cdot l^*} \\ &= e^{2\pi\sqrt{-1}(\sum_{\sigma \in R_H} u^\sigma - \sum_{\sigma \in R_H} u^{\sigma\tau_1}) \cdot l^*} \sum_{\substack{R_H \tau_2 \in R_H \setminus R_G, \\ \sum_{\tau \in R_H} \tau(\tau_1 - \tau_2) = 0}} e^{2\pi\sqrt{-1}\lambda_{\tau_2} \cdot l^*}. \end{aligned} \quad (32)$$

Owing to $M\lambda_\tau \in L$, whether $l^* \in L^* \setminus \mathcal{H}_{G,H}$ belongs to $\Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$ depends only on the class $l^* + ML^*$. \square

There exists a simple condition equivalent to $l^* \in \Gamma_{ext}(G)$:

Corollary 4.2. *For any given crystallographic group G ,*

$$l^* \in L^* \text{ belongs to } \Gamma_{ext}(G) \iff \exists \tau \in R_G \text{ such that } \tau l^* = l^* \text{ and } \nu_\tau \cdot l^* \notin \mathbb{Z}. \quad (33)$$

Proof. Let $R_{G,l^*} \subset R_G$ be the stabilizer subgroup of l^* . Then, $R_L \tau_2 \sim_{l^*} R_L \tau_1$ if and only if $\tau_2 \in \tau_1 R_{G,l^*}$. Hence,

$$\begin{aligned} \sum_{R_L \tau_2 \sim_{l^*} R_L \tau_1} e^{2\pi\sqrt{-1}x^{\{\tau_2 | \nu_{\tau_2}\}} \cdot l^*} &= \sum_{\tau \in R_{G,l^*}} e^{2\pi\sqrt{-1}(x^{\tau_1\tau} + \nu_{\tau_1\tau}) \cdot l^*} \\ &= e^{2\pi\sqrt{-1}(x^{\tau_1} + \nu_{\tau_1}) \cdot l^*} \sum_{\tau \in R_{G,l^*}} e^{2\pi\sqrt{-1}\nu_\tau \cdot l^*}. \end{aligned} \quad (34)$$

The statement follows from the fact that $\tau \mapsto e^{2\pi\sqrt{-1}\nu_\tau \cdot l^*}$ is a homomorphism on R_{G,l^*} . \square

Now that we have defined systematic absences, it is possible to formulate $\Lambda_{ext}(\wp)$ in (10) precisely; for any triplet $(G, H, x + V^{\Sigma_{RH}})$ with $G = R_G \rtimes L$ and $x \in V^H$, we define

$$\Lambda_{ext}(G, H, x + V^{\Sigma_{RH}}) := \left\{ |l^*|^2 : 0 \neq l^* \in L^* \setminus \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \right\}, \quad (35)$$

$$\Lambda_{ext}(G) := \Lambda_{ext}(G, L, V). \quad (36)$$

From the equivalence condition (25), for \wp in (18),

$$\Lambda_{ext}(\wp) := \bigcup_{i=1}^m \Lambda_{ext}(G, H_i, x_i + V^{\Sigma_{RH_i}}). \quad (37)$$

As a result, it is only necessary to consider the case of $\Lambda_{ext}(\wp) = \Lambda_{ext}(G, H, x + V^{\Sigma_{RH}})$ for some $(G, H, x + V^{\Sigma_{RH}})$, so as to discuss the problem (A4) of systematic absences.

5 C-type domains and Conway's topographs

The reduction theory deals with the problem of specifying a domain $D \subset \mathcal{S}_{>0}^N$ satisfying:

(R1) When we put $D[g] := \{gS^t g : S \in D\}$ for any subset $D \subset \mathcal{S}_{>0}^N$ and $g \in GL_N(\mathbb{Z})$, the subgroup of $GL_N(\mathbb{Z})$ consisting of all $g \in GL_N(\mathbb{Z})$ satisfying $D = D[g]$ has only finite elements.

(R2) For any $g \in GL_N(\mathbb{Z})$, D and $D[g]$ do not share interior points, except for $g \in H$.

(R3) $\mathcal{S}_{>0}^N$ is decomposed as follows:

$$\mathcal{S}_{>0}^N = \bigcup_{gH \in GL_N(\mathbb{Z})/H} D[g]. \quad (38)$$

In [6], a topograph was defined from the Selling reduction with $N = 2$. So we shall recall the Selling reduction first; let $A_N := (a_{ij})_{1 \leq i, j \leq N} \in \mathcal{S}_{>0}^N$ be the Gram matrix of the root lattice \mathbb{A}_N having entries as follows:

$$a_{ij} = \begin{cases} 2 & \text{if } i = j, \\ 1 & \text{otherwise.} \end{cases} \quad (39)$$

Using A_N and the inner-product $\langle S, T \rangle := \text{Trace}(ST) = \sum_{i=1}^N \sum_{j=1}^N s_{ij} t_{ij}$ on \mathcal{S}^N , the domain $\mathcal{D}_{Sel}^N \subset \mathcal{S}_{>0}^N$ is defined by

$$\mathcal{D}_{Sel}^N := \{S \in \mathcal{S}_{>0}^N : \langle S, A_N \rangle \leq \langle gS^t g, A_N \rangle \text{ for any } g \in GL_N(\mathbb{Z})\}. \quad (40)$$

If the subgroup of all $g \in GL_N(\mathbb{Z})$ satisfying ${}^t g A_N g = A_N$ is denoted by $H(A_N)$, $\mathcal{S}_{>0}^N$ is partitioned using \mathcal{D}_{Sel}^N by the Selling reduction [27]:

$$\mathcal{S}_{>0}^N = \bigcup_{g \in GL_N(\mathbb{Z})/H(A_N)} \mathcal{D}_{Sel}^N[g], \quad (41)$$

In order to generalize the definition of topographs for general N , we'd like to note that the tessellation of (41) coincides with the one given by Voronoi's two reduction theories [32], [33] in $N = 2, 3$. Hence \mathcal{D}_{Sel}^N ($N = 2, 3$) is same as the *principal domain of the first type* defined in the two reduction theories. In the first reduction theory, the principal domain is defined as the convex cone expanded by $v^t v$ of all the minimal

vectors v of A_N . In the second reduction theory, the same domain is represented as follows:

$$\mathcal{V}(\Phi) := \{S \in \mathcal{S}_{>0}^N : {}^t v S u \leq {}^t u S u \text{ for any } v \in \Phi \text{ and } u \in \mathbb{Z}^N\}, \quad (42)$$

$$\Phi_0^N := \left\{ \pm \sum_{k=1}^N i_k \mathbf{e}_k : i_k = 0, 1 \right\}. \quad (43)$$

Different from \mathcal{D}_{Sel}^N , in the tessellation of $\mathcal{S}_{>0}^N$ given by the first reduction theory, every domain is associated explicitly with the set of integral vectors formed by minimal vectors of a perfect form. Even in the second reduction theory, such an association is provided by using C -type domains (a union of finite L -type domains), instead of L -type domains.

In the following, we choose the association by the second reduction theory, because of Proposition 5.1 on the association of facets of primitive C -type domains and the parallelogram law.

5.1 Topographs for lattices of general rank

In this section, C -type domains and topographs are defined for the general dimension N . We also refer to [24] for more detailed information about C -type domains and its connection with covering problems. In [6], a topograph was defined to explain the Selling reduction with $N = 2$. In contrast, vonorms and conorms were used for the case of $N > 2$.

The idea of vonorms and conorms as invariants of a lattice seems to have originated from the Voronoi vectors defined in Voronoi's second reduction theory; for a fixed $S \in \mathcal{S}_{>0}^N$, if $v \in \mathbb{Z}^N$ satisfies ${}^t v S v = \text{vo}_S(v + 2\mathbb{Z}^N)$, v is called a *Voronoi vector* of S . The *vonorm map* of S is defined as a map from $v + 2\mathbb{Z}^N \in \mathbb{Z}^N / 2\mathbb{Z}^N$ to the representation by the Voronoi vector corresponding to $v + 2\mathbb{Z}^N$.

$$\text{vo}_S(v + 2\mathbb{Z}^N) := \min\{{}^t w S w : w \in v + 2\mathbb{Z}^N\}. \quad (44)$$

Conorms are the Fourier transform of vonorms: when χ is any character on $\mathbb{Z}^N / 2\mathbb{Z}^N$, the *conorm map* of S is defined by:

$$\text{co}_S(\chi) := -\frac{1}{2^{N-1}} \sum_{v+2\mathbb{Z}^N \in \mathbb{Z}^N / 2\mathbb{Z}^N} \text{vo}_S(v + 2\mathbb{Z}^N) \chi(v). \quad (45)$$

In the following, we use vonorm maps in the definition of C -type domains for clarity. Before introducing C -type domains, we shall recall the definition of L -type domains (also called secondary cones *cf.* [30]); The Dirichlet–Voronoi polytope of S is defined by:

$$\text{DV}(S) := \{x \in \mathbb{R}^N : {}^t x S x \leq {}^t (x + l) S (x + l) \text{ for any } l \in \mathbb{Z}^N\}. \quad (46)$$

From the definition, $\text{DV}(S)$ is the intersection of half-spaces:

$$\text{DV}(S) = \bigcap_{0 \neq v \in \mathbb{Z}^N} \{x \in \mathbb{R}^N : {}^t x S x \leq {}^t (x + v) S (x + v)\}. \quad (47)$$

As proved in [33] (*cf.* [7]), $v \in \mathbb{Z}^N$ is a Voronoi vector if and only if the hyperplane $H_{S,v} := \{x \in \mathbb{R}^N : {}^t x S x = {}^t (x + v) S (x + v)\}$ intersects $\text{DV}(S)$.

A tiling of \mathbb{R}^N is given by the Dirichlet–Voronoi polytopes:

$$\mathbb{R}^N = \bigcup_{l \in \mathbb{Z}^N} (\text{DV}(S) + l). \quad (48)$$

The Delone subdivision is a dual tiling of (48). If we let P_S be the set of extreme points of $DV(S)$, and denote the set of all Voronoi vectors v satisfying $p \in H_{S,v}$ by Ψ_p for any $p \in P_S$, then every $v \in \Psi_p$ satisfies ${}^t_p S p = {}^t(p+v)S(p+v)$. Therefore, an ellipsoid $\{x \in \mathbb{R}^N : {}^t(x+p)S(x+p) = {}^t_p S p\}$ passes through all the elements of $\{0\} \cup \Psi_p \subset \mathbb{Z}^N$. If L_p is the convex hull of $\{0\} \cup \Psi_p$, then the Delone subdivision of \mathbb{R}^N is given by:

$$\text{Del}(S) := \bigcup_{p \in P_S, l \in \mathbb{Z}^N} L_p + l. \quad (49)$$

In this case, a *L-type domain* containing S is defined by $\{S_2 \in \mathcal{S}^N : \text{Del}(S) = \text{Del}(S_2)\}$.

For any subset $D \subset \mathcal{S}_{>0}^N$ and $g \in GL_N(\mathbb{Z})$, define $D[g] := \{gS {}^t g : S \in D\}$. Two domains $D_1, D_2 \subset \mathcal{S}_{>0}^N$ are said to be *equivalent* if and only if $D_1[g] = D_2$ holds for some $g \in GL_N(\mathbb{Z})$. Voronoi proved that the number of equivalence classes of *L-type domains* of dimension $\frac{N(N+1)}{2}$ is finite, and provided an algorithm to gain all the equivalence classes [33]. This is the outline of Voronoi's second reduction theory.

For fixed $S \in \mathcal{S}_{>0}^N$, define $\Phi_S := \{[v] : v \in \mathbb{Z}^N \text{ is a Voronoi vector of } S\}$, where $[v]$ represents the class of $v \in \mathbb{Z}^N$ when v and $-v$ are identified. A *C-type domain* containing S is defined by $\mathcal{V}(\Phi) := \{S_2 \in \mathcal{S}_{>0}^N : \Phi_S = \Phi_{S_2}\}$, i.e., a set of elements of $\mathcal{S}_{>0}^N$ having the same Voronoi vectors. Then, $\mathcal{V}(\Phi)$ is a union of finite *L-type domains*, because a set of Voronoi vectors of S is decomposed into $\bigcup_{p \in P_S} \Psi_p$ in different ways, depending on S . From the definition, the Voronoi map is a linear function on $\mathcal{V}(\Phi_S)$ for any $S \in \mathcal{S}_{>0}^N$.

More generally, we shall define *C-type domains* for any set $\Phi := \{[v_1], \dots, [v_n]\}$ of arbitrary size n with $v_1, \dots, v_n \in \mathbb{Z}^N$:

$$\mathcal{V}(\Phi) := \{S \in \mathcal{S}_{>0}^N : {}^t_v S v = \text{vo}_S(v + 2\mathbb{Z}^N) \text{ for any } [v] \in \Phi\}. \quad (50)$$

This is well defined, because both ${}^t_v S v$ and $v + 2\mathbb{Z}^N$ are invariant if v is replaced by $-v$. From the definition, $\mathcal{V}(\Phi)$ is an intersection of the following half-spaces.

$$\mathcal{V}(\Phi) = \bigcap_{[v] \in \Phi} \bigcap_{u+2\mathbb{Z}^N = v+2\mathbb{Z}^N} H^{\geq 0}(u, v), \quad (51)$$

$$H^{\geq 0}(u, v) := \{S \in \mathcal{S}_{>0}^N : {}^t_u S u \geq {}^t_v S v\}. \quad (52)$$

Any $S \in \mathcal{S}_{>0}^N$ is contained in $\mathcal{V}(\Phi_S)$. S is said to be in a *general position* if Φ_S has exactly 2^N elements. If S is in a general position, $\mathcal{V}(\Phi_S)$ includes an open neighbor of S . Such *C-type domains* are called *primitive*. Otherwise, there exist $[v] \neq [u] \in \mathbb{Z}^N$ such that $u + 2\mathbb{Z}^N = v + 2\mathbb{Z}^N$, and S belongs to the following hyperplanes:

$$H(u, v) := \{S \in \mathcal{S}^N : {}^t_u S u = {}^t_v S v\}. \quad (53)$$

Even for such S , by perturbing the entries of S , \tilde{S} in a general position satisfying $\Phi_S \subset \Phi_{\tilde{S}}$ is obtained. As a result, the following partitioning of $\mathcal{S}_{>0}^N$ is obtained.

$$\mathcal{S}_{>0}^N = \bigcup_{\mathcal{V} \in P_N} \mathcal{V}, \quad (54)$$

$$P_N := \{\mathcal{V}(\Phi_S) : S \in \mathcal{S}_{>0}^N, \mathcal{V}(\Phi_S) \text{ is primitive}\}. \quad (55)$$

This tessellation is coarser than that given by *L-type domains*. Hence, similarly with *L-type domains*, the number of equivalence classes of *C-type domains* of dimension $\frac{N(N+1)}{2}$ is finite.

Table 1: Equivalence classes of primitive C -type domains

	Voronoi vectors	Primitive C -type domain
$N = 1$	$\Phi_0^1 := \{[0], [\mathbf{e}_1]\}$	$\mathcal{V}(\Phi_0^1) = \mathcal{S}_{\geq 0}^1$.
$N = 2$	$\Phi_0^2 := \{[0], [\mathbf{e}_1], [\mathbf{e}_2], [\mathbf{e}_1 + \mathbf{e}_2]\}$	$\mathcal{V}(\Phi_0^2) = \bigcap_{1 \leq i < j \leq 3} H^{\geq 0}(\mathbf{e}_i - \mathbf{e}_j, \mathbf{e}_i + \mathbf{e}_j)$
$N = 3$	$\Phi_0^3 := \left\{ \left[\sum_{k=1}^3 i_k \mathbf{e}_k \right] : i_k = 0 \text{ or } 1 \right\}$	$\mathcal{V}(\Phi_0^3) = \bigcap_{1 \leq i < j \leq 4} H^{\geq 0}(\mathbf{e}_i - \mathbf{e}_j, \mathbf{e}_i + \mathbf{e}_j)$
$N = 4$	$\Phi_0^4 := \left\{ \left[\sum_{k=1}^4 i_k \mathbf{e}_k \right] : i_k = 0 \text{ or } 1 \right\}$	$\mathcal{V}(\Phi_0^4) = \bigcap_{1 \leq i < j \leq 5} H^{\geq 0}(\mathbf{e}_i - \mathbf{e}_j, \mathbf{e}_i + \mathbf{e}_j)$
	$\Phi_1^4 := \Phi_0^4 \cup \{[\mathbf{e}_1 - \mathbf{e}_2]\} \setminus \{[\mathbf{e}_1 + \mathbf{e}_2]\}$	$\mathcal{V}(\Phi_1^4) = H^{\geq 0}(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2)$ $\cap \left(\bigcap_{3 \leq i < j \leq 5} H^{\geq 0}(\mathbf{e}_i - \mathbf{e}_j, \mathbf{e}_i + \mathbf{e}_j) \right)$ $\cap \left(\bigcap_{\substack{3 \leq i < j \leq 5, \\ k=1,2}} H^{\geq 0}(\mathbf{e}_i + \mathbf{e}_j + 2\mathbf{e}_k, \mathbf{e}_i + \mathbf{e}_j) \right)$
	$\Phi_2^4 := \Phi_0^4 \cup \{[\mathbf{e}_1 - \mathbf{e}_2], [\mathbf{e}_3 - \mathbf{e}_4]\} \setminus \{[\mathbf{e}_1 + \mathbf{e}_2], [\mathbf{e}_3 + \mathbf{e}_4]\}$	$\mathcal{V}(\Phi_2^4) = H^{\geq 0}(\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2) \cap H^{\geq 0}(\mathbf{e}_3 + \mathbf{e}_4, \mathbf{e}_3 - \mathbf{e}_4)$ $\cap \left(\bigcap_{k=3,4} H^{\geq 0}(\mathbf{e}_1 + \mathbf{e}_2 + 2\mathbf{e}_k, \mathbf{e}_1 - \mathbf{e}_2) \right)$ $\cap \left(\bigcap_{k=1,2} H^{\geq 0}(\mathbf{e}_3 + \mathbf{e}_4 + 2\mathbf{e}_k, \mathbf{e}_3 - \mathbf{e}_4) \right)$ $\cap \left(\bigcap_{i=1}^4 H^{\geq 0}(\mathbf{e}_i - \mathbf{e}_5, \mathbf{e}_i + \mathbf{e}_5) \right)$ $\cap \left(\bigcap_{\substack{i=1,2,k=3,4 \\ \text{or } i=3,4,k=1,2}} H^{\geq 0}(\mathbf{e}_i + \mathbf{e}_5 + 2\mathbf{e}_k, \mathbf{e}_i + \mathbf{e}_5) \right)$

^aHere, we put $\mathbf{e}_{N+1} := -\sum_{i=1}^N \mathbf{e}_i$ in every N -dimensional case.

^bAmong the above domains, only $\mathcal{V}(\Phi_2^4)$ is not an L -type domain, but is a union of two L -type domains. L -type domains for $N = 4$ are listed in [30].

Table 1 lists all the representatives of equivalence classes of C -type domains for $1 \leq N \leq 4$. Each domain $\mathcal{V}(\Phi_i^k)$ in Table 1 has a set of extreme rays provided by:

$$M(\Phi_0^k) := \{v^t v : v = \mathbf{e}_i \ (1 \leq i \leq N), \mathbf{e}_i - \mathbf{e}_j \ (1 \leq i < j \leq N)\}, \quad (56)$$

$$M(\Phi_1^4) := (M(\Phi_0^4) \cup \{D_4\}) \setminus \{v^t v : v = \mathbf{e}_1 - \mathbf{e}_2\}, \quad (57)$$

$$M(\Phi_2^4) := (M(\Phi_0^4) \cup \{v^t v : v = \mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - \mathbf{e}_4\} \cup \{D_4, D_{4,2}\}) \setminus \{v^t v : v = \mathbf{e}_1 - \mathbf{e}_2, \mathbf{e}_3 - \mathbf{e}_4\}, \quad (58)$$

where D_4 and $D_{4,2}$ are the equivalent perfect forms corresponding to the root lattice \mathbb{D}_4 :

$$D_4 := \begin{pmatrix} 2 & 1 & -1 & -1 \\ 1 & 2 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ -1 & -1 & 0 & 2 \end{pmatrix}, \quad D_{4,2} := \begin{pmatrix} 2 & 0 & -1 & -1 \\ 0 & 2 & -1 & -1 \\ -1 & -1 & 2 & 1 \\ -1 & -1 & 1 & 2 \end{pmatrix}. \quad (59)$$

The following proposition claims that every facet of a primitive C -type domain is associated with a set of four vectors satisfying the parallelogram law. Although this is the most important property in our discussion, we could not find references mentioning this explicitly.

Proposition 5.1. *Suppose that two primitive C -type domains $\mathcal{V}(\Phi_{S_1}) \neq \mathcal{V}(\Phi_{S_2})$ have $\binom{N(N+1)}{2} - 1$ -dimensional cone as their intersection. Then $\Phi_{S_1} \cap \Phi_{S_2}$ contains exactly $2^N - 1$ elements. Hence, there exist $u, v \in \mathbb{Z}^N$ such that $\Phi_{S_1} \setminus \Phi_{S_2} = \{[u]\}$ and $\Phi_{S_2} \setminus \Phi_{S_1} = \{[v]\}$. $\mathcal{V}(\Phi_{S_1} \cup \Phi_{S_2}) \subset H(u, v)$ is a common facet of $\mathcal{V}(\Phi_{S_1})$ and $\mathcal{V}(\Phi_{S_2})$. Furthermore, $\{\frac{v+u}{2}, \frac{v-u}{2}\}$ is a primitive set of \mathbb{Z}^N , and $[\frac{v+u}{2}], [\frac{v-u}{2}]$ are elements of $\Phi_{S_1} \cap \Phi_{S_2}$.*

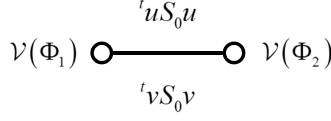


Figure 5: Representations associated with an edge of a topograph.

Proof. From $\mathcal{V}(\Phi_{S_1}) \neq \mathcal{V}(\Phi_{S_2})$, there exist u_1, v_1 such that $u_1 + 2\mathbb{Z}^N = v_1 + 2\mathbb{Z}^N$, $[u_1] \in \Phi_{S_1} \setminus \Phi_{S_2}$, and $[v_1] \in \Phi_{S_2} \setminus \Phi_{S_1}$. Since u_1, v_1 are Voronoi vectors of any $S_3 \in \mathcal{V}(\Phi_{S_1}) \cap \mathcal{V}(\Phi_{S_2})$, we have ${}^t u_1 S_3 l \leq {}^t l S l$ and ${}^t v_1 S_3 l \leq {}^t l S l$. Hence, $\frac{u_1+v_1}{2}$ and $\frac{u_1-v_1}{2}$ are also Voronoi vectors of S_3 . Replacing u_i, v_i with gu_i, gv_i ($g \in GL_N(\mathbb{Z})$) if necessary, we may assume $\frac{u_1-v_1}{2} = \mathbf{e}_1$ and $\frac{u_1+v_1}{2} = m\mathbf{e}_1 + n\mathbf{e}_2$ for some $m, n \in \mathbb{Z}$. Then u_i, v_i and $\frac{u_1 \pm v_1}{2}$ are Voronoi vectors of $\bar{S}_2 := ({}^t \mathbf{e}_i S_3 \mathbf{e}_j)_{1 \leq i, j \leq 2} \in \mathcal{S}_{>0}^2$. Hence S_2 belongs to $\mathcal{V}(\{[\frac{u_1 \pm v_1}{2}], u_1\})$, which is equivalent to $\mathcal{V}(\Phi_0^2)$ in Table 1. As a result, there exists $g \in GL_N(\mathbb{Z})$ such that $g(\frac{u_1-v_1}{2}) = \mathbf{e}_1$ and $g(\frac{u_1+v_1}{2}) = \mathbf{e}_2$. Hence, it is concluded that $\{\frac{u_1+v_1}{2}, \frac{u_1-v_1}{2}\}$ is a primitive set of \mathbb{Z}^N . Suppose that there exists another $(u_2, v_2) \neq (u_1, v_1)$ satisfying $u_2 + 2\mathbb{Z}^N = v_2 + 2\mathbb{Z}^N$, $[u_2] \in \Phi_{S_1} \setminus \Phi_{S_2}$, and $[v_2] \in \Phi_{S_2} \setminus \Phi_{S_1}$. From the dimension of $\mathcal{V}(\Phi_{S_1}) \cap \mathcal{V}(\Phi_{S_2})$, the following must hold for any $S_3 \in \mathcal{S}^N$:

$${}^t(\frac{u_1+v_1}{2})S_3(\frac{u_1-v_1}{2}) = 0 \iff {}^t(\frac{u_2+v_2}{2})S_3(\frac{u_2-v_2}{2}) = 0. \quad (60)$$

We may now assume $\frac{u_1-v_1}{2} = \mathbf{e}_1$ and $\frac{u_1+v_1}{2} = \mathbf{e}_2$. Then, $\{[u_1], [v_1]\} = \{[u_2], [v_2]\}$ is easily obtained. Therefore, $\Phi_{S_1} \setminus \Phi_{S_2}$ and $\Phi_{S_2} \setminus \Phi_{S_1}$ consist of only one element. Now the remaining statements follow immediately. \square

By Proposition 5.1, it is proved that the decomposition (54) is a facet-to-facet tessellation. Hence, (54) is also a face-to-face tessellation by a theorem of Gruber and Ryshkov [13].

Using the V -partitioning (54), a topograph is defined for general N .

Definition 5.1. Define CT_N as the graph which has P_N, E_N as its sets of nodes and edges, respectively.

$$P_N := \{\mathcal{V}(\Phi_S) : S \in \mathcal{S}_{>0}^N, \Phi_S \text{ is primitive}\}, \quad (61)$$

$$E_N := \{e_{\mathcal{V}(\Phi_1), \mathcal{V}(\Phi_2)} : \mathcal{V}(\Phi_1) \neq \mathcal{V}(\Phi_2) \in P_N \text{ share a facet}\}, \quad (62)$$

where $e_{\mathcal{V}(\Phi_1), \mathcal{V}(\Phi_2)}$ connects two nodes $\mathcal{V}(\Phi_1), \mathcal{V}(\Phi_2) \in P_N$. When $S_0 \in \mathcal{S}^N$ is fixed arbitrarily, edges in E_N are associated with two representations of S_0 over \mathbb{Z} by the map:

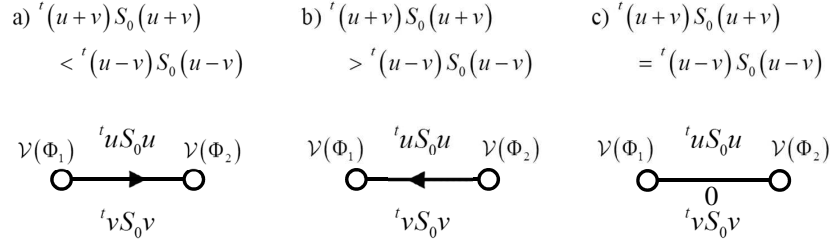
$$f_{S_0}(e_{\mathcal{V}(\Phi_1), \mathcal{V}(\Phi_2)}) := \{{}^t u S_0 u, {}^t v S_0 v\}, \quad (63)$$

where $u, v \in \mathbb{Z}^N$ are taken so that $\{[u+v]\} = \Phi_{S_1} \setminus \Phi_{S_2}$ and $\{[u-v]\} = \Phi_{S_2} \setminus \Phi_{S_1}$. Such an edge is represented in Figure 5. Furthermore, the direction of the edge is defined as in Figure 6. Assuming every edge is oriented by this, we call the pair $CT_{N, S_0} := (CT_N, f_{S_0})$ a topograph of S_0 .

5.2 Topographs for low-dimensional lattices

In this section, the structures of topographs for lattices of rank $N = 2, 3$ are explained. Since the same topic is also discussed in [6], we mention only basic facts necessary in the following sections. In $N = 2, 3$, the structures are also determined from the partitioning (41) of the Selling reduction. In particular, the set of nodes is provided by

$$P_N := \{\mathcal{D}_{Sel}^N[g] = \mathcal{V}({}^t g^{-1} \Phi_0^N) : g \in GL_N(\mathbb{Z})\}. \quad (64)$$



When $\mathcal{V}(\Phi_1), \mathcal{V}(\Phi_2) \in P_N$ have a common facet, there exists $u, v \in \mathbb{Z}^N$ such that $\Phi_1 \setminus \Phi_2 = \{\pm(u+v)\}$, $\Phi_2 \setminus \Phi_1 = \{\pm(u-v)\}$. In this case, the direction of the edge connecting $\mathcal{V}(\Phi_1)$ and $\mathcal{V}(\Phi_2)$ is defined as in [6].

Figure 6: Direction of edges of a topograph CT_{N,S_0} .

By Voronoi's second reduction theory, every node $\mathcal{D}_{Sel}^N[g]$ is associated with ${}^t g^{-1} \Phi_0^N$.

In the following, a lattice L of rank N , a basis b_1, \dots, b_N of L are fixed, and a matrix $(b_1 \dots b_N)$ is denoted by B . In order to clarify ${}^t B B$ is the Gram matrix of (L, B) , we utilize the following notation, instead of f_{tBB} , $CT_{N, {}^t B B}$.

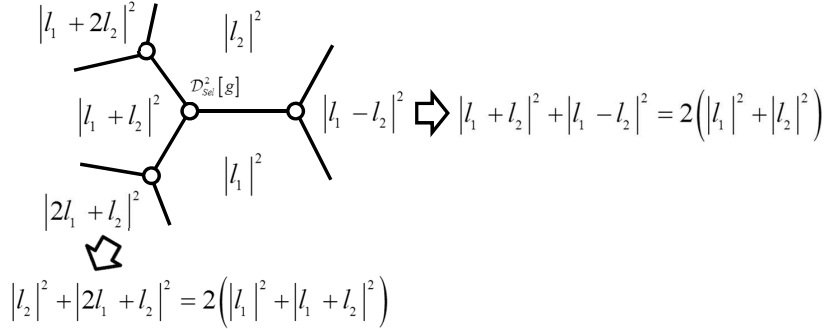
$$f_{(L,B)}(e_{\mathcal{V}(\Phi_1), \mathcal{V}(\Phi_2)}) := f_{tBB}(e_{\mathcal{V}(\Phi_1), \mathcal{V}(\Phi_2)}) = \{|Bu|^2, |Bv|^2\}, \quad (65)$$

$$CT_{N,(L,B)} := CT_{N, {}^t B B}. \quad (66)$$

where $u, v \in \mathbb{Z}^N$ are chosen as in Definition 5.1.

Basic properties of $CT_{2,(L,B)}$ and $CT_{3,(L,B)}$ are explained in the following examples.

Example 1. Case of $CT_{2,(L,B)}$. $\mathcal{D}_{Sel}^2 = \mathcal{V}(\Phi_0^2)$ is a polyhedral cone surrounded by the three hyperplanes in Table 1. Hence, a node of CT_2 is adjacent to three nodes, as in Figure 7.



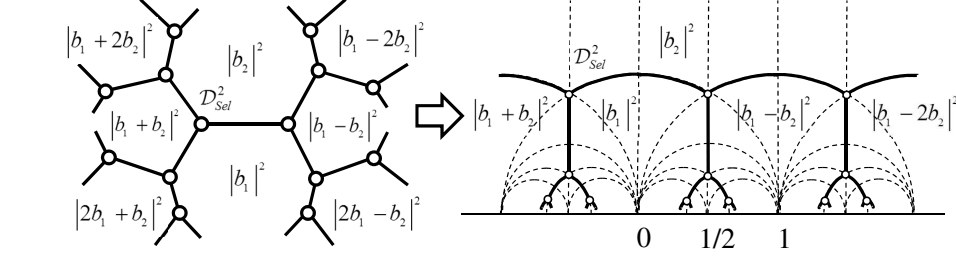
Let $\mathbf{e}_3 := -\mathbf{e}_1 - \mathbf{e}_2$, and $\tau_{ij}^{(2)}$ be the 2×2 matrix satisfying

$$\tau_{ij}^{(2)} \mathbf{e}_i = -\mathbf{e}_j, \quad \tau_{ij}^{(2)} \mathbf{e}_j = \mathbf{e}_i. \quad (67)$$

When we put $(l_1 \ l_2) := (b_1 \ b_2) {}^t g^{-1}$ and $l_3 := -l_1 - l_2$ for the fixed basis b_1, b_2 of L and $g \in GL_2(\mathbb{Z})$, every node $\mathcal{D}_{Sel}^2[g] = \mathcal{V}({}^t g^{-1} \Phi_0^2)$ is an end point of three edges $e_{\mathcal{V}({}^t g^{-1} \Phi_0^2), \mathcal{V}({}^t g^{-1} \tau_{ij}^{(2)} \Phi_0^2)}$ ($1 \leq i < j \leq 3$) that are associated with $\{|l_i|^2, |l_j|^2\}$. (In this figure, any edge between two domains labeled $|k_1|^2$ and $|k_2|^2$ is considered to be associated with $\{|k_1|^2, |k_2|^2\}$. Such a labeling of domains is achieved by embedding the topograph in the upper half plane as in Figure 8, and labeling every domain $D \subset \mathbb{H}$ surrounded by topograph edges with $|l|^2$ of the minimal vector l of some $S \in \iota^{-1}(D)$. This is well-defined, because l is the only minimal vector of all $S \in \iota^{-1}(D)$.)

Figure 7: Local structure of topograph $CT_{2,(L, (b_1 \ b_2))}$.

It can be proved that CT_2 is a tree as in Figure 8. The tree is embedded in $\mathcal{S}_{>0}^2/\mathbb{R}_{>0} \simeq \mathbb{H}$ by mapping a node $\mathcal{V}(t_g^{-1}\Phi_0^2)$ to the perfect form $gA_2^{-1}t_g$, and an edge between $\mathcal{V}(t_{g_1}^{-1}\Phi_0^2)$ and $\mathcal{V}(t_{g_2}^{-1}\Phi_0^2)$ to the geodesic connecting $g_1A_2^{-1}t_{g_1}$, $g_2A_2^{-1}t_{g_2}$.



A one-to-one correspondence between the points of the upper half-plane $\mathbb{H} := \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ and $\mathcal{S}_{>0}^2/\mathbb{R}_{>0}$ is given by the map:

$$\iota : x + \sqrt{-1}y \mapsto \begin{bmatrix} x^2 + y^2 & x \\ x & 1 \end{bmatrix}. \quad (68)$$

Under this identification of \mathbb{H} and $\mathcal{S}_{>0}^2/\mathbb{R}_{>0}$, the action of $g \in GL_2(\mathbb{Z})$ on the latter set by $S \mapsto gS^t g$ coincides with the action of $GL_2(\mathbb{Z})$ on \mathbb{H} by:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \begin{cases} \frac{az+b}{cz+d} & \text{if } ad - bc = 1, \\ \frac{a\bar{z}+b}{c\bar{z}+d} & \text{if } ad - bc = -1, \end{cases} \quad (69)$$

where \bar{z} is the complex conjugate of z .

Figure 8: Embedding of a topograph in the upper half plane.

Example 2. Case of $CT_{3,(L,B)}$. $\mathcal{D}_{Sel}^3 = \mathcal{V}(\Phi_0^3)$ is a polyhedral cone surrounded by the six hyperplanes in Table 1. Hence, a node of CT_3 is adjacent to six nodes, as in Figure 9. All the adjacent nodes of $\mathcal{V}(t_g^{-1}\Phi_0^2)$ are given as $\mathcal{V}(t_g^{-1}\tau_{ij}^{(3)}\Phi_0^2)$ ($1 \leq i < j \leq 4$), where $\mathbf{e}_4 := -\sum_{i=1}^3 \mathbf{e}_i$ and $\tau_{ij}^{(3)}$ is the 3×3 matrix satisfying:

$$\tau_{ij}^{(3)} \mathbf{e}_i = -\mathbf{e}_i, \quad \tau_{ij}^{(3)} \mathbf{e}_j = \mathbf{e}_j, \quad k \neq i, j \implies \tau_{ij}^{(3)} \mathbf{e}_k = \mathbf{e}_i + \mathbf{e}_k. \quad (70)$$

As explained in Figure 9, CT_3 contains two kinds of circuits of length 3 and 6, which correspond to the following fundamental relations of $GL_3(\mathbb{Z})$:

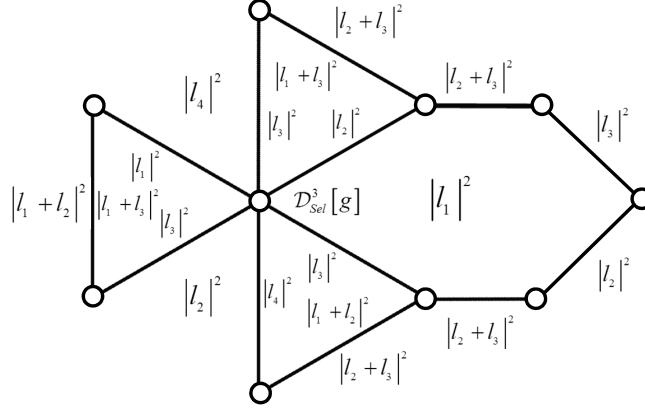
$$\tau_{ij}^{(3)} \tau_{km}^{(3)} \tau_{ij}^{(3)} = \sigma_{ik,jm} \in H(A_3), \quad \left(\tau_{ij}^{(3)} \tau_{ik}^{(3)} \tau_{im}^{(3)} \right)^2 = 1, \quad (71)$$

where i, j, k, m are integers satisfying $\{i, j, k, m\} = \{1, 2, 3, 4\}$, and $\sigma_{ik,jm} \in H(A_3)$ is the 3×3 matrix satisfying:

$$\sigma_{ik,jm} \mathbf{e}_p = -\mathbf{e}_q \text{ for any } (p, q) = (i, k), (k, i), (j, m), (m, j). \quad (72)$$

6 Main results for the distribution rules of systematic absences

We now discuss on the distribution rules of systematic absences on a topograph. Those used in our algorithm are described as theorems, whereas other properties important for powder auto-indexing algorithms are mentioned as facts. As proved in Proposition 4.1, there are only a finite number of types of systematic absences, and they are classified by a triplet $(G, H, x + V^{\sum_{RH}})$.



When we put $(l_1 \ l_2 \ l_3) := (b_1 \ b_2 \ b_3)^t g^{-1}$ and $l_4 := -\sum_{i=1}^3 l_i$ for the fixed basis b_1, b_2, b_3 and $g \in GL_3(\mathbb{Z})$, each node $\mathcal{D}_{Sel}^3[g] = \mathcal{V}({}^t g^{-1} \Phi_0^3)$ is an end point of six edges associated with $\{|l_i|^2, |l_j|^2\}$ ($1 \leq i < j \leq 4$). For any $\{i, j, k, m\} = \{1, 2, 3, 4\}$, two edges associated with $\{|l_i|^2, |l_j|^2\}$, $\{|l_k|^2, |l_m|^2\}$ are contained in a circuit of length 3. On the other hand, two edges associated with $\{|l_i|^2, |l_j|^2\}$, $\{|l_k|^2, |l_l|^2\}$ are contained in a circuit of length 6.

Figure 9: Local structure of a topograph $CT_{3,(L,(b_1 \ b_2 \ b_3))}$.

It is not difficult to prove our theorems if $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$ is contained in $\mathcal{H}_{G,H}$ in (31). (This does not occur in the case of lattices of rank 2.) Otherwise, too many case-by-case considerations are required. Hence, the most difficult part of the theorems is confirmed by direct computation, using the International Tables ($N = 2$) and executing a program ($N = 3$). For confirmation of the case $N = 3$, we verify that the program outputs exactly the same list as the International Tables.

The most important property of $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$ is that L^* is generated by elements of $L^* \setminus \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$, under the assumption that L is the period lattice of the periodic function \wp . Furthermore, if $x + V^{\Sigma_{RH}} \subset V^{H_2}$ holds for some subgroup $H \subset H_2 \subset G$, it is clear that $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) = \Gamma_{ext}(G, H_2, x + V^{\Sigma_{RH_2}})$ holds. Therefore, the following types of systematic absences may be regarded to be *invalid*, and removed from the following consideration.

1. $L^* \setminus \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$ is contained in the reciprocal lattice L_2^* of some $L_2 \supsetneq L$.
2. $x + V^{\Sigma_{RH}} \subset V^{H_2}$ for some subgroup $H \subset H_2 \subset G$.

6.1 Cases of rank 2

According to the International Tables, there are 17 wallpaper groups and 72 types of systematic absences, including invalid ones. By direct computation (or by seeing tables in [14]), it is verified that $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) = \Gamma_{ext}(G, L, V)$ holds in all the valid cases.

In the following, we introduce a short proof of Theorem 1 in the case $H = L$ using a topograph. As a result, Theorem 1 follows for general cases.

Lemma 6.1. *Let L^* be a lattice of rank 2, and $R \subset O(N)$ be the automorphism group of L^* . If $\tau l_1^* = l_1^*$ holds for some $1 \neq \tau \in R$ and primitive vector of $l_1^* \in L^*$, either of the following holds:*

- (a) *there exists $l_2^* \in L^*$ such that $l_1^* \cdot l_2^* = 0$ and l_1^*, l_2^* is a basis of L^* , or*

(b) $\tau^2 = 1$ and there exists $l_2^* \in L^*$ such that $l_1^* = l_2^* + \tau l_2^*$ and $l_2^*, \tau l_2^*$ is a basis of L^* .

Proof. Fix a Gram matrix S_0 of L^* . Let $l_2^* \in L^*$ be the vector satisfying $\{l_1^*, l_2^*\} \in P_2(L^*)$, $l_1^* \cdot l_2^* \geq 0$ and $|l_2^*|^2 = \min\{|l_3^*|^2 : \{l_1^*, l_3^*\} \in P_2(L^*)\}$. The edges of CT_{2,S_0} then have the direction presented in Figure 10. The left figure corresponds to the case (a).

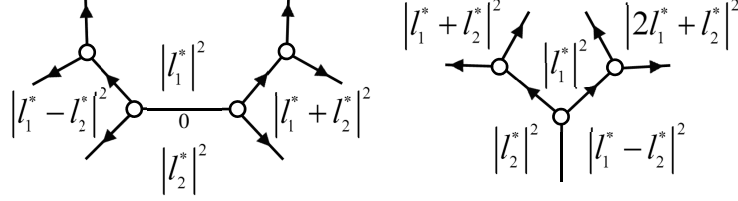


Figure 10: Directions of edges of a topograph associated with $|l_1^*|^2$.

In the case of the right figure, τ fixes the node surrounded by $|l_1^*|^2$, $|l_2^*|^2$, and $|l_1 - l_2|^2$. If we recall that the stabilizer of $\mathcal{V}(\Phi_0^3)$ is given by $H(A_3)$, the action of τ or $-\tau$ on $\{l_1^*, -l_2^*, -l_1^* + l_2^*\}$ coincides with a permutation of $l_1^*, -l_2^*, -l_1^* + l_2^*$. From $\tau \neq 1$ and $\tau l_1^* = l_1^*$, we obtain $\tau l_2^* = l_1^* - l_2^*$. This is equivalent to the case (b). \square

Proof of Theorem 1 in the case of $H = L$. From Corollary 4.2, for any primitive vector $l_1^* \in \Gamma_{ext}(G)$, there exists $\tau \in R$ with $\tau l_1^* = l_1^*$ such that $\nu_\tau \cdot l_1^* \notin \mathbb{Z}$. If (b) of Lemma 6.1 holds, then:

$$\nu_\tau \cdot l_1^* = \nu_\tau^{1+\tau} \cdot l_2^* = \nu_{\tau^2} \cdot l_2^* = \nu_1 \cdot l_2^* \in \mathbb{Z}. \quad (73)$$

Hence, the statement of the theorem follows. \square

6.2 Cases of rank 3

For $N = 3$, we consider the following cases separately:

- (i) $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \cap P_1(L^*)$ is contained in $\mathcal{H}_{G,H}$ in (31);
- (ii) $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \cap P_1(L^*)$ is not contained in $\mathcal{H}_{G,H}$.

Table 2 lists all valid types of systematic absences corresponding to the latter case. For each type, $\Omega \subset L^*/ML^*$ in Corollary 4.1 is given in Table 3.

First, we shall explain some more details of Ito's method, and determine why it does not work appropriately for some types of systematic absences. It was Ito [15] who first proposed using the parallelogram law for powder auto-indexing; if the observed q -values $q_1, q_2, q_3, q_4 \in \Lambda^{obs}$ satisfy $2(q_1 + q_2) = q_3 + q_4$, the method assumes that there exist $l_1^*, l_2^* \in L^*$ satisfying the following:

$$q_1 = |l_1^*|^2, \quad q_2 = |l_2^*|^2, \quad q_3 = |l_1^* + l_2^*|^2, \quad q_4 = |l_1^* - l_2^*|^2. \quad (74)$$

If (74) is true, the Gram matrix of the sublattice expanded by l_1^*, l_2^* is determined. In order to obtain candidate solutions for L^* , it is necessary to construct a 3×3 Gram matrix from combinations of these sublattices. In order to simplify the combination procedure, it is very desirable that $\{l_1^*, l_2^*\}$ in (74) is a primitive set of L^* . Otherwise, it is necessary to check whether L^* has rank 3 sublattices that are more plausible as a solution once L^* is obtained. The program of Visser [31] which adopted Ito's method also implicitly requires $\{l_1^*, l_2^*\} \in P_2(L^*)$ [10].

However, according to the following fact, in some types of systematic absences, $\{l_1^*, l_2^*\} \subset L^*$ is never a primitive set of L^* , if $\{l_1^*, l_2^*\}$ satisfies (74).

Table 2: Types of systematic absences having $\Gamma_{ext}(G, H, xl_1 + yl_2 + zl_3 + V^{\Sigma_{R_H}}) \not\subset \mathcal{H}_{G,H}$.

Space group G (No. ^a)	R_H	$(x, y, z)^b$						
Type-A (Face-centered lattice)			Type-B (Body-centered lattice)			Type-F (Body-centered)		
$F d d 2$ (43)	C_2	$(0, 0, z)$	$I 4_1/a$ (88)	C_i	$(0, \frac{1}{4}, \frac{1}{8})$	$P \bar{4} 3 n$ (218)	C_2	$(x, 0, \frac{1}{2})$
$F d d d$ (70)	C_2	$(x, 0, 0)$	$I 4_1/a$ (88)	C_i	$(\frac{1}{4}, 0, \frac{1}{8})$	$P m \bar{3} n$ (223)	C_2	$(\frac{1}{4}, y, y + \frac{1}{2})$
$F d d d$ (70)	D_2	$(0, 0, 0)$	$I 4_1/a m d$ (141)	C_{2h}	$(0, \frac{1}{4}, \frac{1}{8})$	$P m \bar{3} n$ (223)	C_{2v}	$(x, \frac{1}{2}, 0)$
$F d d d$ (70)	D_2	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$I 4_1/a m d$ (141)	C_{2h}	$(0, \frac{1}{4}, \frac{1}{8})$	$P m \bar{3} n$ (223)	C_{2v}	$(x, 0, \frac{1}{2})$
$F d \bar{3}$ (203)	C_2	$(x, 0, 0)$	Type-C			Type-F (Body-centered)		
$F d \bar{3}$ (203)	T	$(0, 0, 0)$	$I 4_1/a m d$ (141)	C_2	$(x, \frac{1}{4}, \frac{1}{8})$	$I \bar{4} 3 d$ (220)	C_3	(x, x, x)
$F d \bar{3}$ (203)	T	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$I 4_1/a c d$ (142)	C_2	$(\frac{1}{4}, y, \frac{1}{8})$	$I a \bar{3} d$ (230)	C_3	(x, x, x)
$F 4_1 3 2$ (210)	C_2	$(x, 0, 0)$	Type-D			Type-G		
$F 4_1 3 2$ (210)	T	$(0, 0, 0)$	$P 3 1 c$ (159)	C_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 4_2 3 2$ (208)	D_2	$(\frac{1}{4}, 0, \frac{1}{2})$
$F 4_1 3 2$ (210)	T	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$P \bar{3} 1 c$ (163)	C_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 4_2 3 2$ (208)	D_2	$(\frac{1}{4}, \frac{1}{2}, 0)$
$F d \bar{3} m$ (227)	C_{2v}	$(x, 0, 0)$	$P \bar{3} 1 c$ (163)	D_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P \bar{4} 3 n$ (218)	S_4	$(\frac{1}{4}, 0, \frac{1}{2})$
$F d \bar{3} m$ (227)	T_d	$(0, 0, 0)$	$P \bar{3} 1 c$ (163)	D_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P \bar{4} 3 n$ (218)	S_4	$(\frac{1}{4}, \frac{1}{2}, 0)$
$F d \bar{3} m$ (227)	T_d	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$P 6_3$ (173)	C_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P m \bar{3} n$ (223)	D_{2d}	$(\frac{1}{4}, 0, \frac{1}{2})$
Type-A (Body-centered lattice)			$P 6_3/m$ (176)	C_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P m \bar{3} n$ (223)	D_{2d}	$(\frac{1}{4}, \frac{1}{2}, 0)$
$I 4_1$ (80)	C_2	$(0, 0, z)$	$P 6_3/m$ (176)	C_{3h}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	Type-H		
$I 4_1/a$ (88)	C_2	$(0, 0, z)$	$P 6_3/m$ (176)	C_{3h}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 4_3 3 2$ (212)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
$I 4_1/a$ (88)	S_4	$(0, 0, 0)$	$P 6_3 2 2$ (182)	C_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 4_3 3 2$ (212)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
$I 4_1/a$ (88)	S_4	$(0, 0, \frac{1}{2})$	$P 6_3 2 2$ (182)	D_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 4_1 3 2$ (213)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
$I 4_1 2 2$ (98)	C_2	$(0, 0, z)$	$P 6_3 2 2$ (182)	D_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 4_1 3 2$ (213)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
$I 4_1 2 2$ (98)	D_2	$(0, 0, 0)$	$P 6_3 m c$ (186)	C_{3v}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	Type-I		
$I 4_1 2 2$ (98)	D_2	$(0, 0, \frac{1}{2})$	$P \bar{6} 2 c$ (190)	C_3	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$I 4_1 3 2$ (214)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
$I 4_1 m d$ (109)	C_{2v}	$(0, 0, z)$	$P \bar{6} 2 c$ (190)	C_{3h}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$I 4_1 3 2$ (214)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$
$I \bar{4} 2 d$ (122)	C_2	$(0, 0, z)$	$P \bar{6} 2 c$ (190)	C_{3h}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	Type-J		
$I \bar{4} 2 d$ (122)	S_4	$(0, 0, 0)$	$P 6_3/m m c$ (194)	C_{3v}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$I 4_1 3 2$ (214)	D_2	$(\frac{1}{4}, 0, \frac{1}{4})$
$I \bar{4} 2 d$ (122)	S_4	$(0, 0, \frac{1}{2})$	$P 6_3/m m c$ (194)	D_{3h}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$I 4_1 3 2$ (214)	D_2	$(\frac{1}{4}, 0, \frac{1}{4})$
$I 4_1/a m d$ (141)	C_2	$(x, x, 0)$	$P 6_3/m m c$ (194)	D_{3h}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$I \bar{4} 3 d$ (220)	S_4	$(\frac{1}{4}, 0, \frac{1}{4})$
$I 4_1/a m d$ (141)	C_{2v}	$(0, 0, z)$	Type-E			$I \bar{4} 3 d$ (220)	S_4	$(\frac{1}{4}, 0, \frac{1}{4})$
$I 4_1/a m d$ (141)	D_{2d}	$(0, 0, 0)$	$P 6_2$ (171)	C_2	$(\frac{1}{2}, \frac{1}{2}, z)$	Type-K		
$I 4_1/a m d$ (141)	D_{2d}	$(0, 0, \frac{1}{2})$	$P 6_4$ (172)	C_2	$(\frac{1}{2}, \frac{1}{2}, z)$	$I 4_1 3 2$ (214)	C_2	$(x, 0, \frac{1}{4})$
$I 4_1/a c d$ (142)	C_2	$(x, x, \frac{1}{4})$	$P 6_2 2 2$ (180)	C_2	$(\frac{1}{2}, 0, z)$	$I \bar{4} 3 d$ (220)	C_2	$(x, 0, \frac{1}{4})$
Type-B (Face-centered lattice)			$P 6_2 2 2$ (180)	D_2	$(\frac{1}{2}, 0, 0)$	$I a \bar{3} d$ (230)	C_2	$(\frac{1}{8}, y, -y + \frac{1}{4})$
$F d d d$ (70)	C_i	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$P 6_2 2 2$ (180)	D_2	$(\frac{1}{2}, 0, \frac{1}{2})$	Type-L		
$F d d d$ (70)	C_i	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$P 6_4 2 2$ (181)	C_2	$(\frac{1}{2}, 0, z)$	$I a \bar{3} d$ (230)	C_2	$(x, 0, \frac{1}{4})$
$F d \bar{3}$ (203)	C_{3i}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 6_4 2 2$ (181)	D_2	$(\frac{1}{2}, 0, 0)$	Type-M		
$F d \bar{3}$ (203)	C_{3i}	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$P 6_4 2 2$ (181)	D_2	$(\frac{1}{2}, 0, \frac{1}{2})$	$I a \bar{3} d$ (230)	D_2	$(\frac{1}{8}, 0, \frac{1}{8})$
$F 4_1 3 2$ (210)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	Type-F (Primitive)			$I a \bar{3} d$ (230)	S_4	$(\frac{1}{8}, 0, \frac{1}{8})$
$F 4_1 3 2$ (210)	D_3	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	$P 4_2 3 2$ (208)	C_2	$(x, \frac{1}{2}, 0)$	Type-N		
$F d \bar{3} m$ (227)	D_{3d}	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	$P 4_2 3 2$ (208)	C_2	$(x, 0, \frac{1}{2})$	$I a \bar{3} d$ (230)	D_3	$(\frac{1}{8}, \frac{1}{8}, \frac{1}{8})$
$F d \bar{3} m$ (227)	D_{3d}	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	$P \bar{4} 3 n$ (218)	C_2	$(x, \frac{1}{2}, 0)$			

^aNumber assigned to every space group in [14].

^bThe fractional coordinate (x, y, z) indicates the point $xl_1 + yl_2 + zl_3 \bmod L_2 \in V$ ($x, y, z \in [0, 1)$), where $L_2 \subset L$ is the Bravais lattice, and l_1, l_2, l_3 is the basis of L_2 chosen by the method in "Notation and symbols".

Table 3: Necessary and sufficient condition^a for $\sum_{i=1}^3 u_i l_i^* \notin \mathcal{H}_{G,H}$ to belong to $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$.

Type	Necessary and sufficient condition		
A	face-centered	\implies	$u_1 + u_2 + u_3 \equiv 2 \pmod{4}$
	body-centered	\implies	$2u_2 + u_3 \equiv 2 \pmod{4}$
B	face-centered	\implies	$\{u_1, u_2, u_3\} \equiv \{0, 0, 2\}, \{0, 2, 2\} \pmod{4}$
	body-centered	\implies	$\{u_1, u_2, u_3\} \equiv \{1, 1, 0\}, \{3, 3, 0\}, \{1, 3, 2\} \pmod{4}$
C	$(u_1, u_2, u_3) \equiv (1, 1, 0) \pmod{2}$.		
D	$u_1 \equiv u_2 \pmod{3}, u_3 \equiv 1 \pmod{2}$.		
E	$u_1, u_2 \equiv 0 \pmod{2}, u_3 \not\equiv 0 \pmod{3}$.		
F	primitive	\implies	$\{u_1, u_2, u_3\} \equiv \{1, 1, 1\} \pmod{2}$
	body-centered	\implies	$\{u_1, u_2, u_3\} \equiv \{0, 0, 2\}, \{2, 2, 2\} \pmod{4}$
G	$\{u_1, u_2, u_3\}$	$\begin{cases} \equiv \{0, 0, 1\}, \{1, 1, 1\} \pmod{2}, \text{ and} \\ \not\equiv \{0, 1, 2\}, \{0, 2, 3\} \pmod{4}. \end{cases}$	
H	$\{u_1, u_2, u_3\}$	$\begin{cases} \equiv \{0, 0, 0\}, \{0, 0, 1\} \pmod{2}, \text{ and} \\ \not\equiv \{0, 1, 2\}, \{0, 2, 3\}, \{0, 0, 0\}, \{2, 2, 2\} \pmod{4}. \end{cases}$	
I	$\{u_1, u_2, u_3\} \equiv \{0, 0, 2\}, \{0, 2, 2\} \pmod{4}$.		
J	$\{u_1, u_2, u_3\}$	$\begin{cases} \not\equiv \{0, 0, 0\}, \{0, 2, 2\}, \{1, 1, 2\}, \{1, 2, 3\}, \{2, 3, 3\} \pmod{4}, \text{ and} \\ \not\equiv \{0, 2, 4\}, \{0, 4, 6\}, \{0, 1, 1\}, \{1, 1, 4\}, \{0, 3, 3\}, \{3, 3, 4\}, \{0, 1, 7\}, \{1, 4, 7\}, \{0, 3, 5\}, \{3, 4, 5\} \pmod{8}. \end{cases}$	
K	$\{u_1, u_2, u_3\} \equiv \{2, 2, 2\} \pmod{4}$.		
L	$\{u_1, u_2, u_3\} \equiv \{0, 1, 1\}, \{0, 1, 3\}, \{0, 3, 3\}, \{2, 2, 2\} \pmod{4}$.		
M	$\{u_1, u_2, u_3\}$	$\begin{cases} \not\equiv \{0, 0, 0\}, \{0, 2, 2\}, \{1, 1, 2\}, \{1, 2, 3\}, \{2, 3, 3\} \pmod{4}, \text{ and} \\ \not\equiv \{0, 2, 4\}, \{0, 4, 6\} \pmod{8}. \end{cases}$	
N	$\{u_1, u_2, u_3\} \not\equiv \{0, 0, 0\}, \{1, 1, 2\}, \{1, 2, 3\}, \{2, 3, 3\} \pmod{4}$.		

^a $\langle l_1^*, l_2^*, l_3^* \rangle$ is the reciprocal basis of the basis l_1, l_2, l_3 of \mathbb{R}^3 taken as in the footnote of Table 2.

Fact 1. If $(G, H, x + V^{\Sigma_{RH}})$ is a type B or N, the following set is empty:

$$\left\{ \{l_1^*, l_2^*\} \in P_2(L^*) : l_1^*, l_2^*, l_1^* + l_2^*, l_1^* - l_2^* \notin \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \right\}. \quad (75)$$

To remove adverse effects of systematic absences from Ito's algorithm, it has been proposed that a formula other than the parallelogram law should be used [9]:

$$|l_1^* + ml_2^*|^2 - |l_1^* - ml_2^*|^2 = m(|l_1^* + l_2^*|^2 - |l_1^* - l_2^*|^2). \quad (76)$$

However, it has not been ascertained whether the equation works appropriately for all types of systematic absences. As another example, the following was proposed to obtain a rank 3 solution directly.

$$|l_1^*|^2 + |l_2^*|^2 + |l_3^*|^2 + |l_1^* + l_2^* + l_3^*|^2 = |l_1^* + l_2^*|^2 + |l_1^* + l_3^*|^2 + |l_2^* + l_3^*|^2. \quad (77)$$

The above equation has a similar property to the parallelogram law:

Fact 2. If $(G, H, x + V^{\Sigma_{RH}})$ is of type B, C, F, G, or N, the following set is empty:

$$\left\{ \{l_1^*, l_2^*, l_3^*\} \in P_3(L^*) : l_1^*, l_2^*, l_3^*, l_1^* + l_2^*, l_1^* + l_3^*, l_2^* + l_3^*, l_1^* + l_2^* + l_3^* \notin \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \right\}. \quad (78)$$

The following two theorems are intended to resolve any this kind of problems caused by systematic absences. There, it is proved that there exist “infinitely many” elements of $P_2(L^*)$ or $P_3(L^*)$ satisfying some common properties, and connected subgraphs of a topograph containing infinitely many nodes are formed from the lengths of lattice vectors belonging to such elements. Existence of infinitely many elements is required not to fail to gain the true solution, because only a finite subset $\Lambda^{obs} \subset \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$ is available in computation. From the latter claim, it is guaranteed that rather large

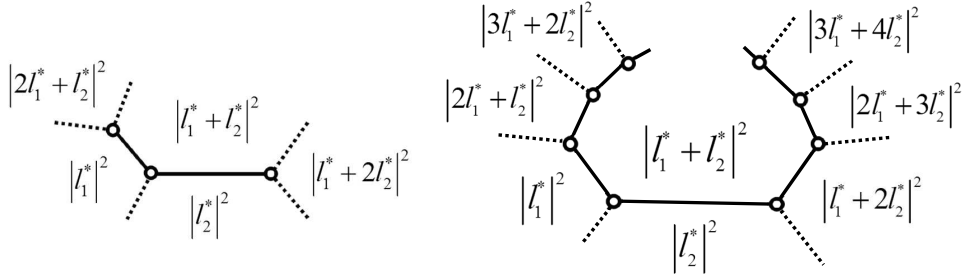


Figure 11: A subgraph of a topograph corresponding to the equation $3|l_1^*|^2 + |l_1^* + 2l_2^*|^2 = |2l_1^* + l_2^*|^2 + 3|l_2^*|^2$ and a chain formed from such subgraphs.

subgraph is formed for the true solution, even from finite elements belonging to Λ^{obs} . This is useful to select better candidate solutions and reduce computation time.

We shall explain the meaning of Theorem 3 before describing the statements; in short, it claims that the equation $3|l_1^*|^2 + |l_1^* + 2l_2^*|^2 = |2l_1^* + l_2^*|^2 + 3|l_2^*|^2$ works appropriately regardless of the type of systematic absences. When (L^*, B^*) represents the reciprocal lattice of the lattice to obtain as a solution and a matrix whose columns are a basis of L^* , this new equation corresponds to the subgraph of $CT_{2,(L^*, B^*)}$ consisting of three nodes and two edges in the left-hand of Figure 11. A subgraph of $CT_{2,(L^*, B^*)}$ containing infinitely many nodes as the right-hand of Figure 11 is formed by linking such a subgraph obtained from $q_r, q_t, q_s, q_u \in \Lambda_{ext}(G, H, x + V^{\Sigma_{RH}})$ satisfying $3q_r + q_t = 3q_s + q_u$.

Theorem 3. *For any type of systematic absences $(G, H, x + V^{\Sigma_{RH}})$ with $G = R_G \ltimes L$, the following subset of $P_2(L^*)$ includes infinitely many elements.*

$$\tilde{P}_2(L^*) := \left\{ \{l_1^*, l_2^*\} \in P_2(L^*) : ml_1^* + (m-1)l_2^* \notin \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \text{ for any } m \in \mathbb{Z} \right\}. \quad (79)$$

More precisely, for any open convex cone $U \subset \mathbb{R}^N$ satisfying $U \cap \mathcal{H}_{G,H} = \emptyset$, the following holds:

- (1) *There exists $\{l_1^*, l_2^*\} \in P_2(L^*)$ such that $\{ml_1^* + nl_2^* : m, n \in \mathbb{Z}_{\geq 0}\} \subset U$. If $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \subset \mathcal{H}_{G,H}$, such $\{l_1^*, l_2^*\} \in P_2(L^*)$ belongs to $\tilde{P}_2(L^*)$.*
- (2) *Assume that $(G, H, x + V^{\Sigma_{RH}})$ is one of the types in Table 2, and $\{l_1^*, l_2^*\} \in P_2(L^*)$ satisfy*
 - (a) $\{l_1^*, l_2^*\} \subset U$, and
 - (b) $l_1^*, l_2^*, 2l_1^* + l_2^*, l_1^* + 2l_2^*, 2l_1^* + 3l_2^*, 3l_1^* + 2l_2^*, 3l_1^* + 4l_2^*$ (i.e., all the vectors other than $l_1^* + l_2^*$ in the right-hand of Figure 11) do not belong to $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$.

Then, $\{l_1^, l_2^*\}$ is an element of $\tilde{P}_2(L^*)$. Furthermore, $\{l_1^*, l_2^*\} \in P_2(L^*)$ satisfying the two properties exists.*

Proof. If $\tilde{P}_2(L^*)$ is not empty, it includes infinitely many elements because, for any $n \in \mathbb{Z}$, $\{-nl_1^* - (n-1)l_2^*, (n+1)l_1^* + nl_2^*\} \in P_2(L^*)$ belongs to $\tilde{P}_2(L^*)$ if and only if $\{l_1^*, l_2^*\} \in \tilde{P}_2(L^*)$. Hence, the first statement follows immediately from (1) and (2).

(1) is almost straightforward. In particular, the existence of $\{l_1^*, l_2^*\} \in P_2(L^*)$ satisfying $\{l_1^*, l_2^*\} \subset U$ is proved in Lemma 6.2. Thus, $ml_1^* + nl_2^* \in U$ holds for any integers $m, n \geq 0$. Therefore, if $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}}) \subset \mathcal{H}_{G,H}$, $\pm(ml_1^* + nl_2^*) \notin \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$ follows. As a result, we obtain $\{l_1^*, l_2^*\} \in \tilde{P}_2(L^*)$.

In order to prove (2), let M be the order of R_G . We first define the following sets:

$$\Gamma_{ext,M}(G, H, x + V^{\Sigma_{R_H}}) := \left\{ l^* + ML^* : l^* \in \Gamma(G, H, x + V^{\Sigma_{R_H}}) \setminus \mathcal{H}_{G,H} \right\}, \quad (80)$$

$$\tilde{P}_{2,M}(L^*) := \left\{ \{l_1^* + ML^*, l_2^* + ML^*\} : \begin{array}{l} \{l_1^*, l_2^*\} \in P_2(L^*) \text{ such that for any } m \in \mathbb{Z}, \\ ml_1^* + (m-1)l_2^* + ML^* \notin \Gamma_{ext,M}(G, H, x + V^{\Sigma_{R_H}}) \end{array} \right\}. \quad (81)$$

Note that $\Gamma_{ext,M}(G, H, x + V^{\Sigma_{R_H}})$ is equal to Ω in Corollary 4.1. By direct computation, it is verified that $\tilde{P}_{2,M}(L^*)$ equals the following set:

$$\left\{ \{l_1^* + ML^*, l_2^* + ML^*\} : \begin{array}{l} \{l_1^*, l_2^*\} \in P_2(L^*) \text{ satisfying for some } m = k, \dots, k+6 \ (k \in \mathbb{Z}), \\ ml_1^* + (m-1)l_2^* \notin \Gamma_{ext,M}(G, H, x + V^{\Sigma_{R_H}}) \end{array} \right\}. \quad (82)$$

Consequently, if $\{l_1^*, l_2^*\} \in P_2(L^*)$ with $\{l_1^*, l_2^*\} \subset U$ satisfy the assumption of (2), $\{l_1^* \bmod M, l_2^* \bmod M\} \in \tilde{P}_{2,M}(L^*)$, and therefore $\{l_1^*, l_2^*\} \in \tilde{P}_2(L^*)$ is obtained from the definition of U . Furthermore, $\tilde{P}_{2,M}(L^*) \neq \emptyset$ holds for any type of systematic absences presented in Table 4. Therefore, from Lemma 6.2, there exists $\{l_1^*, l_2^*\} \in P_2(L^*)$ such that $\{l_1^* + ML^*, l_2^* + ML^*\} \in \tilde{P}_{2,M}(L^*)$ and $\{l_1^*, l_2^*\} \subset U$. In this case, $\{l_1^*, l_2^*\} \in P_2(L^*)$ also holds. \square

The following lemma is used in the proof of Theorem 3. Although it is straightforward, we provide a proof.

Lemma 6.2. *Let $L \subset \mathbb{R}^N$ be a lattice of rank N , and $1 \leq m \leq N$ be an integer. Then, any open cone $U \subset \mathbb{R}^N$ contains a primitive set $\{l_1, \dots, l_m\} \in P_m(L)$. Furthermore, when $M > 0$ is a positive integer and $\{k_1, \dots, k_m\} \in P_m(L)$, U contains infinitely many $\{l_1, \dots, l_m\} \in P_m(L)$ satisfying $l_i - k_i \in ML$ for any $1 \leq i \leq m$.*

Proof. We prove the first statement by induction. Since $\{al : a \in \mathbb{Q}, l \in L\} = \{al : a \in \mathbb{Q}, l \in P_1(L)\}$ is dense in \mathbb{R}^N , there exist $a \in \mathbb{Q}$ and $l \in P_1(L)$ such that $al \in U$. Hence, $l \in U$ is obtained. Next suppose that $m < N$ and there is $T \in P_m$ with $T \subset U$. Then there exists $l \in L$ such that $T \cup \{l\} \in P_{m+1}$. Furthermore, for any $l_2 \in T$, there is $\epsilon > 0$ such that $U_2 := \{x \in \mathbb{R}^N : (1 - \epsilon)|x|^2 |l_2|^2 \leq (x \cdot l_2)^2\}$ is contained in U . Clearly there is an integer $s > 0$ such that $l + sl_2 \in U_2$. As a result, $T \cup \{l + sl_2\} \in P_{m+1}$ is a subset of U . In order to prove the second statement, we fix a basis $l_1, \dots, l_N \in U$ of L . Then there exists $g \in GL(\mathbb{Z}/M\mathbb{Z})$ such that $k_i = gl_i$ for any $1 \leq i \leq m$. When the subgroup of $g \in GL(\mathbb{Z})$ with positive entries is denoted by $GL_+(\mathbb{Z})$, the natural map $GL_+(\mathbb{Z}) \rightarrow G := \{g \in GL(\mathbb{Z}/M\mathbb{Z}) : \det g = \pm 1 \bmod M\}$ is an epimorphism. Let $g_0 \in GL_+(\mathbb{Z})$ be an element belonging to the inverse image of $g \bmod M$. Then $g_0 l_1, \dots, g_0 l_N$ are all contained in U and satisfy $g_0 l_i - k_i \in ML$ ($1 \leq i \leq m$). \square

Theorem 4 claims that a solution of rank 3 is obtained from the combination of two subgraphs of a topograph for lattices expanded by l_1^* and l_i^* ($i = 2, 3$) as in the right-hand of Figure 11 and $|\pm l_1^* + l_2^* + l_3^*|^2 \in \Lambda_{ext}(G, H, x + V^{\Sigma_{R_H}})$. By using such subgraphs with infinitely many edges for the purpose, the “sort criterion for zones” proposed in Section 7.3 is provided a theoretical foundation.

Theorem 4. *Using the same notation as Theorem 3, let $\tilde{P}_3(L^*)$ be the set of $\{l_1^*, l_2^*, l_3^*\} \in P_3(L^*)$ satisfying (a) and (b) with both $i = 2, 3$.*

- (a) $\pm l_1^* + l_2^* + l_3^* \notin \Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$.
- (b) $ml_1^* + (m-1)(-l_1^* + l_i^*) \notin \Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$ for any $m \in \mathbb{Z}$, or $ml_i^* + (m-1)(l_1^* - l_i^*) \notin \Gamma_{ext}(G, H, x + V^{\Sigma_{R_H}})$ for any $m \in \mathbb{Z}_{\geq 0}$.

Table 4: Density of $L^* \setminus \Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$ in L^* .

Type	$\tilde{P}_{2,M}(L^*)$ in $P_{2,M}(L^*)$	$\tilde{P}_{3,M}(L^*)$ in $P_{3,M}(L^*)$
A	0.321	0.286
B	0.286	0.143
C	0.476	0.190
D	0.341	0.209
E	0.736	0.604
F	0.714	0.571
G	0.214	0.027
H	0.429	0.058
I	0.714	0.571
J	0.071	0.022
K	0.857	0.786
L	0.107	0.004
M	0.036	0.004
N	0.036	0.004

^a M is the cardinality of R_G . The densities are computed by dividing the cardinality of $\tilde{P}_{i,M}(L^*)$ by that of $P_{i,M}(L^*)$ ($i = 2, 3$), where $P_{i,M}$ is defined by

$$P_{i,M}(L^*) := \{(l_1^* + ML^*, \dots, l_k^* + ML^*) : \{l_1^*, \dots, l_k^*\} \in P_i(L^*)\}. \quad (83)$$

^bThe densities of G, H, J, L, M, N, which consist of only primitive and body-centered cubic lattices, are rather small. We were not able to find another combination of q -values with larger densities. As a practical measure against small densities, the parallelogram law is used in the enumeration procedure of Table 6, in addition to $3|l_1^*|^2 + |l_1^* + 2l_2^*|^2 = |2l_1^* + l_2^*|^2 + 3|l_2^*|^2$. (This is effective except for the type N, owing to Fact 1.) Furthermore, as explained in Section 7.3, condition (b) in Theorem 4 is relaxed in the actual algorithm, and subgraphs with relatively many edges are prioritized.

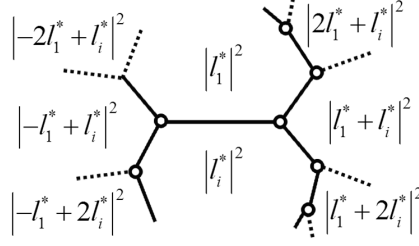


Figure 12: Connected subgraph of a topograph consisting of edges associated with $\{\{l_i, ml_1^* + (m-1)(-l_1 + l_i^*) : m \in \mathbb{Z}\} \text{ or } \{\{l_1, ml_i^* + (m-1)(l_1^* - l_i^*) : m \in \mathbb{Z}_{\geq 0}\}\}$.

$\tilde{P}_3(L^*)$ then contains infinitely many elements, regardless of the type of systematic absences.

Remark 2. This theorem may be considered to refer the vectors associated with the edges in the circuit of length 6 in Figure 9; the 6 edges are associated with one of the four vectors in the following parallelogram law:

- (a) $2(|l_1|^2 + |l_2 + l_3|^2) = |l_1 + l_2 + l_3|^2 + |-l_1 + l_2 + l_3|^2$,
- (b) $2(|l_1|^2 + |l_i|^2) = |l_1 + l_i|^2 + |-l_1 + l_i|^2$ ($i = 2$ or 3).

Proof. By Lemma 6.2, any open convex cone $U \subset \mathbb{R}^N \setminus \mathcal{H}_{G,H}$ contains $\{l_1^*, -l_1^* + l_2^*, -l_1^* + l_3^*\} \in P_3(L^*)$. Hence, U includes $\pm l_1^* + l_2^* + l_3^*, ml_1^* + (m-1)(-l_1 + l_i^*)$ ($m \in \mathbb{Z}$) and $ml_i^* + (m-1)(l_1^* - l_i^*)$ ($m \in \mathbb{Z}_{\geq 0}$). Consequently, if $\Gamma_{ext}(G, H, x + V^{\Sigma_{RH}})$ is contained in $\mathcal{H}_{G,H}$, the statement is obtained immediately.

If $(G, H, x + V^{\Sigma_{RH}})$ is one of the types in Table 2, let M be the order of R_G , and $\Gamma_{ext,M}(G, H, x + V^{\Sigma_{RH}})$ and $\tilde{P}_{2,M}(L^*)$ be the sets defined in the proof of Theorem 3. Furthermore, we define

$$\tilde{P}_{3,M}(L^*) := \left\{ \begin{array}{l} \{l_1^* + ML^*, l_2^* + ML^*, l_3^* + ML^*\} : \begin{array}{l} \{l_1^*, l_2^*, l_3^*\} \in P_3(L^*), \\ \pm l_1^* + l_2^* + l_3^* + ML^* \in \Gamma_{ext,M}(G, H, x + V^{\Sigma_{RH}}), \\ \{l_1^* + ML^*, -l_1^* - l_i^* + ML^*\} \text{ or } \{l_i^* + ML^*, -l_1^* - l_i^* + ML^*\} \\ \text{belongs to } \tilde{P}_{2,M}(L^*) \text{ (} 2 \leq i \leq 3 \text{)} \end{array} \end{array} \right\}.$$

By direct calculation, it is verified that $\tilde{P}_{3,M}(L^*) \neq \emptyset$, regardless of the type of systematic absences (See Table 4). From Lemma 6.2, there exist infinitely many $\{l_1^*, -l_1^* + l_2^*, -l_1^* + l_3^*\} \in P_3(L^*)$ such that $\{l_1^* + ML^*, -l_1^* + l_2^* + ML^*, -l_1^* + l_3^* + ML^*\} \in \tilde{P}_{3,M}(L^*)$ and $\{l_1^*, -l_1^* + l_2^*, -l_1^* + l_3^*\} \subset U$. In this case, $\{l_1^*, l_2^*, l_3^*\} \in \tilde{P}_3(L^*)$ also holds. \square

7 Algorithm for powder auto-indexing

In order to elaborate our new algorithm, we first define a data structure that is assumed to be implemented in the program. The set Λ^{obs} of observed lattice vector lengths is input as an array $\langle (q_i[0], q_i[1]) : 1 \leq i \leq M \rangle$ consisting of pairs of a q -value $q_i[0]$ and its estimated error $q_i[1] = \text{Err}[q_i[0]]$. In our algorithm, every candidate for a Gram matrix of L^* is provided as a matrix with entries $\sum_{i=1}^M n_i q_i[0]$ having coefficients $n_i \in \frac{1}{4}\mathbb{Z}$. At various stages of powder auto-indexing, the propagated errors of the entries are useful for making statistical judgments and strengthening the algorithm against observation errors in the q -values. For this purpose, a data structure of a formal sum of elements of Λ^{obs} with coefficients $n_i \in \frac{1}{4N}\mathbb{Z}$ is implemented in Conograph (N is a fixed positive

Table 5: Procedures for enumeration of four q -values satisfying the parallelogram law.

void enumerateItoSolutions (Λ^{obs}, c, Ans)		
(Input)	Λ^{obs}	: array of N_{peak} pairs of a q -value $q_i[0]$ and its approximated error $q_i[1]$.
	$c > 0$: parameter setting error tolerance level.
(Output)	Ans	: array of a sequence $(\{q_r, q_s\}, \{q_t, q_u\})$, where q_r, q_s, q_t, q_u are elements of Λ^{obs} satisfying

$$\begin{cases} |\text{Val}(2q_r + 2q_s - q_t - q_u)| \leq c \min\{2\text{Err}(q_r + q_s), \text{Err}(q_t + q_u)\} & (\text{Parallelogram law}), \\ \left(\sqrt{q_r[0]} - \sqrt{q_s[0]}\right)^2 \leq q_t[0], q_u[0] \leq \left(\sqrt{q_r[0]} + \sqrt{q_s[0]}\right)^2 & (\text{Positive-definite condition}). \end{cases}$$

```

1:  (Start)  Set a sorted sequence  $S := \langle q_i + q_j : 1 \leq i \leq j \leq N_{peak} \rangle$  of formal sums.
2:          for  $i := 1$  to  $\frac{1}{2}N_{peak}(N_{peak} + 1)$  do
3:              Let  $1 \leq J_{min}, J_{max} \leq N_{peak}$  be integers satisfying
4:               $J_{min} \leq j \leq J_{max} \iff |\text{Val}(2S[i] - S[j])| \leq 2c\text{Err}(S[i])$ .
5:              for  $j := J_{min}$  to  $J_{max}$  do
6:                  if  $|\text{Val}(2S[i] - S[j])| \leq c\text{Err}(S[j])$  then
7:                       $\{q_r, q_s\} := \text{getTerms}(S[i])$ ,
8:                       $\{q_t, q_u\} := \text{getTerms}(S[j])$ .
9:                      if  $(\sqrt{q_r[0]} - \sqrt{q_s[0]})^2 \leq q_t[0], q_u[0] \leq (\sqrt{q_r[0]} + \sqrt{q_s[0]})^2$  then
10:                         insert  $(\{q_r, q_s\}, \{q_t, q_u\})$  in  $Ans$ .
11:                     end if
12:                 end if
13:             end for
14:         end for

```

integer). The data structure is equipped with the order $<$ and functions getTerms , Val , Err as follows:

$$\sum_{i=1}^M a_i q_i < \sum_{i=1}^M b_i q_i \stackrel{\text{def}}{\iff} \sum_{i=1}^M a_i q_i[0] < \sum_{i=1}^M b_i q_i[0], \quad (85)$$

$$\text{getTerms}\left(\sum_{i=1}^M n_i q_i\right) := \{q_i : 1 \leq i \leq M, n_i \neq 0\}, \quad (86)$$

$$\text{Val}\left(\sum_{i=1}^M n_i q_i\right) := \sum_{i=1}^M n_i q_i[0], \quad (87)$$

$$\text{Err}\left(\sum_{i=1}^M n_i q_i\right) := \left(\sum_{i=1}^M n_i^2 (q_i[1])^2\right)^{1/2}. \quad (88)$$

In particular, if Val and Err are called with the argument $\sum_{i=1}^M n_i q_i$, they return the value and the propagated error of $\sum_{i=1}^M n_i q_i[0]$, respectively.

7.1 Algorithm for $N = 2$

In this case, according to Theorem 1, the method using the parallelogram law works sufficiently.

On output of the procedure in Table 5, each entry $(\{q_r, q_s\}, \{q_t, q_u\})$ in Ans satisfies the parallelogram law $2(q_r + q_s) = q_t + q_u$, and corresponds to the 2×2 positive-definite

symmetric matrix:

$$(\{q_r, q_s\}, \{q_t, q_u\}) \mapsto \begin{pmatrix} \text{Val}(q_r) & \frac{1}{2}\text{Val}(q_t - q_r - q_s) \\ \frac{1}{2}\text{Val}(q_t - q_r - q_s) & \text{Val}(q_s) \end{pmatrix}. \quad (89)$$

Here we used the assumption that there exist $l_1^*, l_2^* \in L^*$ such that $q_r = |l_1^*|^2$, $q_s = |l_2^*|^2$, $q_t = |l_1^* + l_2^*|^2$.

7.2 Algorithm for $N = 3$

The theorems in Section 6.2 state how to construct the Gram matrices of lattices of rank 3 from elements of Λ^{obs} satisfying the equation $3|l_1^*|^2 + |l_1^* + 2l_2^*|^2 = |2l_1^* + l_2^*|^2 + 3|l_2^*|^2$. Considering the case in which powder diffraction patterns contain only a small number of peaks, it is better to also use q -values satisfying the parallelogram law in the enumeration algorithm. Hence, the following two computational assumptions are used in the algorithm of Table 6:

- (C1) If $q_r, q_s, q_t, q_u \in \Lambda^{obs}$ satisfy $2(q_r + q_s) = q_t + q_u$, then there are $l_1^*, l_2^* \in L^*$ such that $q_r = |l_1^*|^2$, $q_s = |l_2^*|^2$, $q_t = |l_1^* + l_2^*|^2$, $q_u = |l_1^* - l_2^*|^2$.
- (C2) If $q_r, q_s, q_t, q_u \in \Lambda^{obs}$ satisfy $3q_r + q_t = 3q_s + q_u$, then there are $l_1^*, l_2^* \in L^*$ such that $q_r = |l_1^*|^2$, $q_s = |l_2^*|^2$, $q_t = |l_1^* + 2l_2^*|^2$, $q_u = |2l_1^* + l_2^*|^2$.

The computation time of the procedure in Table 6 is roughly proportional to N_{zone}^2 , where N_{zone} is the size of A_2 immediately after (2). This is estimated as follows. The size of A_3 in (3) is approximately $4N_{zone}$, if the cases $Q_1 = Q_2$ or $Q_3 = Q_4$ are ignored. When N_{peak} is the cardinality of Λ^{obs} , the average number of $(R_1, R_2, R_3, R_4) \in A_2$ satisfying (a) or (b) with regard to fixed $(Q_1, Q_2, Q_3, Q_4) \in A_2$ is approximated as $\frac{4N_{zone}}{N_{peak}}$. Hence, the number of combinations of (Q_1, Q_2, Q_3, Q_4) , $(R_1, R_2, R_3, R_4) \in R_2$ and $q_k \in \Lambda^{obs}$ is roughly equal to $4N_{zone} \cdot \frac{4N_{zone}}{N_{peak}} \cdot N_{peak} = 16N_{zone}^2$. Steps (1)–(3) take much less time than (4). Hence it is concluded that the time is proportional to N_{zone}^2 .

The enumeration is completed by calling the function in Table 6, after which the following procedures are required before outputting solutions.

1. Transform every enumerated Gram matrix into a Niggli-reduced form [22] (which is standard in crystallography).
2. Bravais lattice determination.
3. Sort solutions by some figures of merit (*e.g.*, de Wolff figure of merit, widely used in powder auto-indexing [11]).
4. Remove duplicate solutions.

7.3 Speed-up method using topographs

As described in Section 7.2, the computation time of the algorithm in Table 6 is proportional to the square of N_{zone} in A_2 immediately after (2). Thus, an effective way to speed up the algorithm is to reduce the size of A_2 .

When reducing the size of A_2 , it is necessary to retain elements obtained from $q_r, q_s, q_t, q_u \in \Lambda^{obs}$ for which the assumption (C1) or (C2) is true, because they are essential to obtain the true solution. A new criterion to sort the elements of A_2 is defined for the purpose.

Every entry $(\{R_1, R_2\}, \{R_3, R_4\})$ of A_2 consists of four formal sums satisfying the parallelogram law $2(R_1 + R_2) = R_3 + R_4$. Hence, it corresponds to the subgraph

Table 6: Enumeration algorithm for three-dimensional lattices constructed from q -values.

void enumerateThreeDimLattices ($\Lambda^{obs}, c, \det S_{min}, \det S_{max}, Ans$)		
(Input)	Λ^{obs}, c	: same as in Table 5.
	$\det S_{min}, \det S_{max}$: threshold on minimum and maximum determinants of output matrices.
(Output)	Ans	: array of 3×3 positive-definite symmetric matrices.
(1) By the method in Table 5, enumerate $(\{q_r, q_s\}, \{q_t, q_u\})$ of Λ^{obs} satisfying the parallelogram law and insert in A_2 . (Here, A_2 is an array of four formal sums $(\{Q_1, Q_2\}, \{Q_3, Q_4\})$.)		
(2) Enumerate four elements $(\{q_r, q_s\}, \{q_t, q_u\})$ of Λ^{obs} satisfying the equation $3q_r + q_t = 3q_s + q_u$. This is done by a similar method as in Table 5. Using a new formal sum $q_{-1} := \frac{-2q_r + q_s + q_u}{2} = \frac{q_r - 2q_s + q_t}{2}$, two sets of q -values satisfying the parallelogram law are generated from such $(\{q_r, q_s\}, \{q_t, q_u\})$:		

$$2(q_{-1} + q_r) = q_s + q_u, \quad 2(q_{-1} + q_s) = q_r + q_t. \quad (90)$$

Check whether A_2 contains $(\{q_w, q_r\}, \{q_s, q_u\})$ or $(\{q_w, q_s\}, \{q_r, q_t\})$ for some $1 \leq w \leq N_{peak}$. If this is false (this suggests that $q_{-1} := \frac{-2q_r + q_s + q_u}{2} = \frac{q_r - 2q_s + q_t}{2}$ is undetected in Λ^{obs} owing to some observational reason), insert $(\{\frac{-2q_r + q_s + q_u}{2}, q_r\}, \{q_s, q_u\})$ and $(\{\frac{q_r - 2q_s + q_t}{2}, q_s\}, \{q_r, q_t\})$ in A_2 .

- (3) for every entry $(\{Q_1, Q_2\}, \{Q_3, Q_4\}) \in A_2$, insert (Q_1, Q_2, Q_3, Q_4) , (Q_1, Q_2, Q_4, Q_3) , (Q_2, Q_1, Q_3, Q_4) , (Q_2, Q_1, Q_4, Q_3) in a new array A_3 .
- (4) For every $(Q_1, Q_2, Q_3, Q_4) \in A_3$, search $(R_1, R_2, R_3, R_4) \in A_3$ satisfying either of the following:
 - (a) $Q_1 = R_1 \in \Lambda^{obs}$.
 - (b) $Q_1, R_1 \notin \Lambda^{obs}$ and $|\text{Val}(Q_1 - R_1)| \leq c\text{Err}(Q_1 - R_1)$.

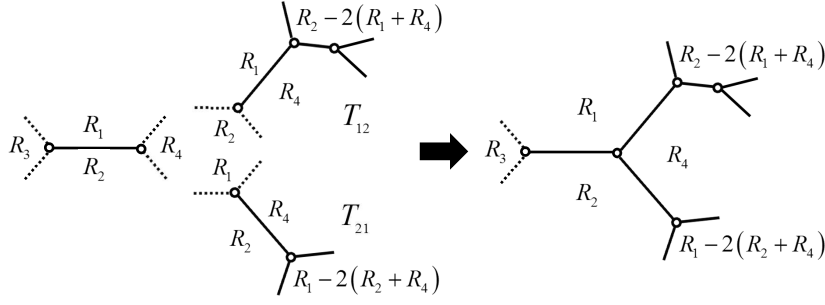
In addition, for every $q_k \in \Lambda^{obs}$, assume that there exists l_1^*, l_2^*, l_3^* satisfying^a

$$\begin{aligned} Q_1 &\approx R_1 = |l_1|^2, \quad Q_2 = |l_2|^2, \quad Q_3 = |l_1 + l_2|^2, \\ R_2 &= |l_3|^2, \quad R_3 = |l_1 + l_3|^2, \quad q_k = |l_1 + l_2 + l_3|^2. \end{aligned} \quad (91)$$

Then, the Gram matrix $S := (l_i \cdot l_j)_{1 \leq i, j \leq 3}$ is obtained as the following 3×3 symmetric matrix:

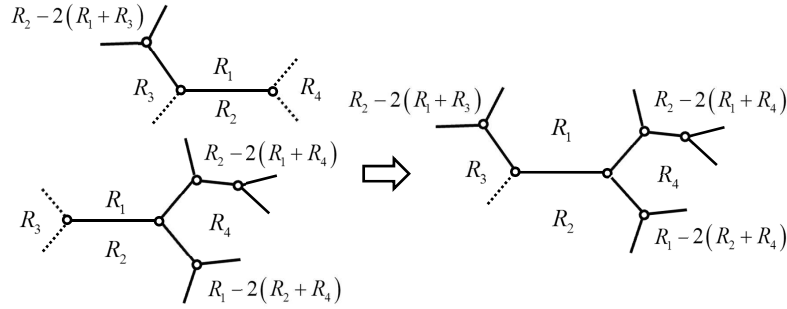
$$\begin{pmatrix} Q_1 & \frac{1}{2}(Q_3 - Q_1 - Q_2) & \frac{1}{2}(R_3 - Q_1 - R_2) \\ \frac{1}{2}(Q_3 - Q_1 - Q_2) & Q_2 & \frac{1}{2}(Q_1 - Q_3 - R_3 + q_k) \\ \frac{1}{2}(R_3 - Q_1 - R_2) & \frac{1}{2}(Q_1 - Q_3 - R_3 + q_k) & R_2 \end{pmatrix}, \quad (92)$$

using Val and Err, the values and propagated errors of the entries are computable. If $\det S_{min} \leq \det S \leq \det S_{max}$, insert S in Ans .



^aEvery three graphs on the left-hand side have a common node, which is an end point of three edges associated with $\{R_1, R_2\}$, $\{R_1, R_4\}$, and $\{R_2, R_4\}$. This figure illustrates how these graphs are unified.

Figure 13: Extension of a topograph (1/2).



^aThe left-hand two subgraphs contain the subgraph in Figure ?? commonly. They are unified as above.

Figure 14: Extension of a topograph (2/2).

of a topograph as in Figure 2. Table 7 explains how a subgraph T of a topograph is constructed from these substructures. The elements of A_2 utilized to obtain the subgraph are output in \tilde{A}_2 .

In the procedure of Table 7, the subgraph T is expanded to only one side. The whole subgraph composed of $e := (\{R_1, R_2\}, \{R_3, R_4\}) \in A_2$ and entries of A_2 is obtained by calling the recursive procedure twice, setting $\tilde{e} := (\{R_1, R_2\}, R_3, R_4)$ and $(\{R_1, R_2\}, R_4, R_3)$ respectively in the second argument. Let $C(e)$ be the cardinality of the union of the two \tilde{A}_2 output by calling the recursive procedure twice as above. This $C(e)$ is considered to quantify the size of the subgraph finally obtained.

If either (C1) or (C2) is wrong for $e \in A_2$, $C(e)$ will result in a small number, because the new q -values required to extend a new edge (*i.e.*, $q \in \Lambda^{obs}$ in line 2 of Table 7) would rarely be found in Λ^{obs} . This suggests $C(e)$ provides an effective sort criterion for elements of A_2 . Part (b) of Theorem 3 indicates that this criterion remains effective under the influence of systematic absences.

Now any $e := (\{R_1, R_2\}, \{R_3, R_4\}) \in A_2$ corresponds to a 2×2 Gram matrix $S(e) \in \mathcal{S}^2$ by the following map:

$$S((\{R_1, R_2\}, \{R_3, R_4\})) \mapsto \begin{pmatrix} \text{Val}(R_1) & \frac{1}{2} \text{Val}(R_3 - R_1 - R_2) \\ \frac{1}{2} \text{Val}(R_3 - R_1 - R_2) & \text{Val}(R_2) \end{pmatrix}. \quad (93)$$

When $C(e)$ is used as a sort criterion, $e \in A_2$ with smaller $\det S(e)$ is prioritized as a result, because $\Lambda^{obs} \subset [q_{min}, q_{max}]$ contains a smaller number of ${}^t u S(e) u$ ($u \in \mathbb{Z}^2$) if $\det S(e)$ has a larger value. This property of $C(e)$ is also desirable, because (4) in Table 6 (and Theorem 4) requires $\{l_1^*, l_i^*\} \in P_2(L^*)$ ($i = 2, 3$) to obtain the true solution L^* .

Table 7: Recursive procedure to form a subgraph of a topograph using \tilde{e} and elements of A_2 .

void expandSubtopograph($A_2, \tilde{e}, \tilde{A}_2, T$)	
(Input)	A_2 : array of four formal sums $(\{Q_1, Q_2\}, \{Q_3, Q_4\})$ with $Q_i := \sum_{j=1}^{N_{peak}} n_{ij} q_j$. Furthermore, it is assumed that every entry satisfies $\text{Val}(2Q_1 + 2Q_2 - Q_3 - Q_4) \leq c \min\{2\text{Err}(Q_1 + Q_2), \text{Err}(Q_3 + Q_4)\},$ for some fixed constant c , and at least three of Q_i ($1 \leq i \leq 4$) belong to Λ^{obs} (<i>i.e.</i> , there is $q \in \Lambda^{obs}$ such that $Q_i = q$).
(Output)	\tilde{e} : $(\{R_1, R_2\}, R_3, R_4)$ satisfying $(\{R_1, R_2\}, \{R_3, R_4\}) \in A_2$. \tilde{A}_2 : subset of A_2 . T : subgraph ^a of a topograph composed of substructures corresponding to $(\{Q_1, Q_2\}, \{Q_3, Q_4\}) \in \tilde{A}_2$. (If such subgraphs are not unique, \tilde{A}_2 containing larger number of entries is prioritized.)
1:	(Start) For $(i, j) = (1, 2), (2, 1)$, let S_{ij} be the set defined by
2:	$S_{ij} := \begin{cases} \{(\{R_i, R_4\}, R_j, q) \in A_2 : q \in \Lambda^{obs}\} & \text{if } R_j \in \Lambda^{obs}, \\ \emptyset & \text{otherwise.} \end{cases}$
3:	$\tilde{T}_{12} := \emptyset$. $\tilde{A}_{12} := \emptyset$. $\tilde{T}_{21} := \emptyset$. $\tilde{A}_{21} := \emptyset$.
4:	for $i = 1$ to 2 do
5:	Take $j \in \{1, 2\} \setminus \{i\}$.
6:	for $\tilde{e}_2 \in S_{ij}$ do
7:	Call expandSubtopograph($A_2, \tilde{e}_2, T_{ij}, A_{ij}$).
8:	if $ \tilde{A}_{ij} < A_{ij} $ then
9:	$\tilde{T}_{ij} := T_{ij}$.
10:	$\tilde{A}_{ij} := A_{ij}$.
11:	end if
12:	end for
13:	end for
14:	Set $\tilde{A}_2 := \{(\{R_1, R_2\}, \{R_3, R_4\})\} \cup \tilde{A}_{12} \cup \tilde{A}_{21}$.
15:	Construct T by unifying $\tilde{T}_{12}, \tilde{T}_{21}$ and $(\{R_1, R_2\}, \{R_3, R_4\})$, as in Figure 13.

^aIn the actual computation, it is not necessary to construct a subgraph T , because a sort criterion for elements of A_2 is defined using only \tilde{A}_2 . However, as T has a data structure of binary trees consisting of finite nodes and edges, it is not difficult to implement T in a program.

Using $C(e)$, a sublattice L_i^* expanded by $\{l_1^*, l_i^*\} \in P_2(L^*)$ is prioritized over any other sublattices of L_i^* of rank 2.

To summarize the above discussion, it is only necessary to insert the following procedures between (2) and (3) in Table 6, in order to reduce the computation time.

- (2i) For each element $e := (\{R_1, R_2\}, \{R_3, R_4\}) \in A_2$, compute $C(e)$ by calling the procedure in Table 6. (In this case, any $e_2 \in \tilde{A}_2$ satisfies $C(e) = C(e_2)$. Hence, the number of calls of the procedure is less than the size of A_2 .)
- (2ii) Sort A_2 in descending order of $C(e)$. For $e_1, e_2 \in A_2$ satisfying $C(e_1) = C(e_2)$, they may be sorted by $e_1 < e_2 \xLeftrightarrow{def} \det S(e_1) < \det S(e_2)$.
- (2iii) Remove the $(N_{zone} + 1)$ th-to-last elements of A_2 using a fixed integer $N_{zone} > 0$. (The default value of N_{zone} used in Conograph is given in (97) of section 8.1. When the default value is used, the computation time is approximately proportional to N_{peak}^4 .)

In Section 8.2, we shall see how the computation time is decreased in practice by this improvement.

7.4 Problems with the quality of powder diffraction patterns

Among the problems described in (A1)–(A6) and ($\tilde{A}3$) of Section 3, we have not yet clarified how to handle ($\tilde{A}3$). In this section, we explain how missing and false elements in Λ^{obs} influence the powder auto-indexing results. This issue is related to the quality of powder diffraction patterns as observational data.

Using a peak search program equipped with Conograph, it is not difficult to obtain the positions of all the peak heights above a given threshold automatically and rather uniformly. Although the ability to decompose overlapping peaks depends on the peak search software, this is not a great problem in our algorithm, because an almost identical solution will be obtained even if $q_1 \in \Lambda^{obs}$ is replaced with q_2 very close to q_1 (to some degree, $\text{Err}[q_1]$ and $\text{Err}[q_2]$ will absorb the difference between q_1 and q_2).

By the algorithm in Table 6, normally, multiple Gram matrices of the reciprocal lattice L^* of the correct solution L are generated from Λ^{obs} . This is because a lattice L^* has as infinitely many Gram matrices as bases of L^* . Note that Gram matrices of different bases are computed from different q -values basically. Therefore when these Gram matrices are transformed into a reduced form, they are a bit different matrices owing to observational errors, but rather close to each other.

In order to reduce the influence of $\epsilon_1 > 0$ of ($\tilde{A}3$), the existence of such duplicate solutions is very useful. Suppose that m Gram matrices of L^* are generated from the set of q -values $\Lambda_{ext}(\wp) \cap [q_{min}, q_{max}]$ containing no observational errors by applying our enumeration method. Then, when the same method is applied to Λ^{obs} containing observational errors, the enumeration process fails to obtain the correct solution, only when none of approximate solutions of the m Gram matrices them are generated. The failure rate becomes very small as m increases, regardless of the magnitude of ϵ_1 . If the size of Λ^{obs} is augmented, or equivalently the range $[q_{min}, q_{max}]$ is magnified, m increases naturally. As a result, the enumeration success rate increases monotonically as the size of Λ^{obs} increases. However, it should be noted that the time for enumeration of solutions is roughly proportional to the fourth power of the size of Λ^{obs} , if the default parameters of Conograph are used. In addition, the success probability will not increase much once q_{max} reaches some observational limit, because larger q -values have larger errors owing to peak overlap and observational accuracy.

Next, we discuss the influence of ϵ_2 in ($\tilde{A}3$). In this case, the success rate in obtaining the true solution remains same if Λ^{obs} is replaced with $\Lambda^{obs} \cup \{q^{obs}\}$ which contains a false element q^{obs} . (Of course, such q^{obs} increases the time for enumeration.)

In conclusion, the enumeration process is considered to be robust to missing and false elements in Λ^{obs} , although such elements increase the computation time. Indeed, the procedure to sort candidate solutions executed after the enumeration is more sensitive to missing and false elements, because figures of merit are severely affected by these elements. In general, for a poor quality powder diffraction pattern, it is very difficult to judge which solution is correct, even if the q -values of respective solutions are manually compared to actual peak-positions (see Example 6 in Section 8.2).

Example 5 shows the case of a two-phase sample. Both lattice parameters are acquired by Conograph. However it seems to be almost impossible to judge which is the correct second phase parameter in this case.

8 Implementation and results of Conograph

Before introducing the input parameters and results of Conograph, we explain the circumstances in which the program was verified. The Conograph source code is written in C++ and OpenMP, and compiled with Mingw (GNU Compiler Collection for Windows). The computer used for the test has an Intel i7-2620 (2.70 GHz) processor and 12 GB RAM. The processor can execute parallel computing with eight hyper-threads, all of which were used during the test. Additionally, we confirmed that a computer with 4 GB RAM was able to carry out the same test.

8.1 Input parameters of Conograph

The input parameters used in the enumeration process are listed in Table 8. In particular, “AUTO” is used to set parameters that depend on respective powder diffraction patterns. The parameters required after the enumeration are not given here; they will be introduced at another time.

Table 8: Default parameters of Conograph.

Symbol	Meaning	Default
$\Delta 2\theta$	Zero-point shift (degree)	0.
c	Tolerance level for errors in q -values	
	Characteristic X-rays or neutron reactor sources	1.5
	Synchrotron X-rays or neutron spallation sources	1.
N_{peak}	Number of q -values used	AUTO
N_{zone}	Threshold for the maximum number of $(\{Q_1, Q_2\}, \{Q_3, Q_4\})$	AUTO
N_{sol}	Threshold for the maximum number of solutions	
	First trial	AUTO
	When the first trial failed because of too many solutions ^a	$64000 \times m$ ($m \in \mathbb{Z}_{>0}$)
Vol_{min}	Threshold for the minimum volume of \mathbb{R}^3/L (\AA^3)	AUTO
Vol_{max}	Threshold for the maximum volume of \mathbb{R}^3/L (\AA^3)	AUTO

^a N_{sol} is set to 64000 at most when AUTO is used, in order to prevent memory allocation errors in memoryless computers. When the number of enumerated solutions exceeds N_{sol} , Conograph removes those with a smaller unit-cell volume. As a result, the true solution is sometimes also removed. The simplest method of resolving this is to reduce Vol_{max} or increase N_{sol} when unsatisfactory results are obtained. This time, we chose the latter in order to measure computation time. However, such a trial-and-error approach is not necessary, because an alternative method is adopted in Conograph. This will be introduced at another time.

In the following, we provide an explanation of the parameters and formulas needed to compute AUTO.

- (1) **Zero-point shift $\Delta 2\theta$.** Depending on the type of diffractometers, x -coordinates of a powder diffraction pattern is given as the following function of q -values.

$$2\theta = 2\sin^{-1}\frac{\lambda\sqrt{q}}{2} + \Delta 2\theta, \quad (94)$$

$$\text{Time-of-flight} = \sum_{i=1}^n c_i q^{-i/2}. \quad (95)$$

Using these equations, x -coordinates are transformed to q -values before executing powder auto-indexing. Among λ , $\Delta 2\theta$ and c_i ($1 \leq i \leq n$), only the value of $\Delta 2\theta$ is unknown. Conograph users are recommended to set $\Delta 2\theta = 0$ normally, because successful results were obtained even in cases of very large zero-point shift, such as $\Delta 2\theta \approx 0.2$ degree. After auto-indexing, it is possible to refine the lattice parameters and zero-point shift simultaneously using a nonlinear least-squares method.

- (2) **Tolerance level c for errors in q -values.** Table 5 provides a usage example of c . By setting a larger value of c , a wider range of combinations of q -values is searched.
- (3) **Number of q -values used.** This parameter determines the number of q -values in Λ^{obs} . After sorting the q -values in Λ^{obs} into ascending order, the $(N_{peak} + 1)$ th-to-last parameters are removed before the powder auto-indexing commences. As explained in Section 7.4, both the computation time and success rate increase as N_{peak} is magnified. The default value of N_{peak} is calculated by:

$$N_{peak} := \min\{\#\{q < 10/d^2 : q \in \Lambda^{obs}\}, 48\}, \quad (96)$$

where $\#T$ is the cardinality of the set T , and d is the minimum threshold for the distance between two lattice points. In Conograph, d is set to 2.0\AA .

In (96), $N_{peak} \leq 48$ is forced, because it is frequently meaningless to use large q -values owing to severe peak overlap. However, 48 is still much larger than the 20–30 that are normally adopted for powder auto-indexing. Conograph uses such many q -values for another reason; in (A5) and the paragraphs following (A6), we explained that $q_{max} > D_N := 3 \cdot 2^{-1/3} d^{-2} \approx 0.6\text{\AA}^{-2}$ is required at least theoretically. (In (96), $q_{max} > 4D_N$ is adopted.) Figure 15 presents the range of the first N_{peak} q -values in the test data of Table 15. Owing to the threshold of 48, the interval $[q_{min}, q_{max}]$ is frequently much smaller than 0.6\AA^{-2} , although powder auto-indexing succeeded in all the test data. We suppose 48 q -values might be insufficient in some exceptional cases. Users are recommended to increase N_{peak} manually, if results obtained with the default parameters are unsatisfactory.

- (4) **Threshold for the maximum size of $(\{Q_1, Q_2\}, \{Q_3, Q_4\})$.**

$$N_{zone} := \frac{1}{3}N_{peak}(N_{peak} + 1). \quad (97)$$

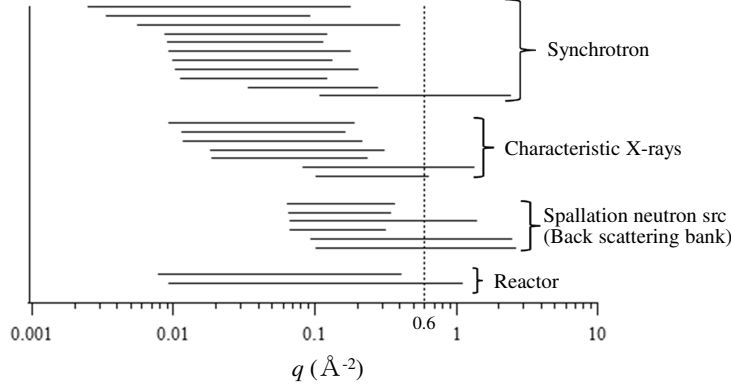
- (5) **Threshold for the maximum number of candidate solutions.**

$$N_{sol} := \min\{64000, N_{zone}^2\}. \quad (98)$$

- (6) **Threshold for the minimum and maximum of the volume of \mathbb{R}^3/L .**

$$\text{Vol}_{min} := \max\{5, v_{20}^{-1}\}, \quad (99)$$

$$\text{Vol}_{max} := 30\text{Vol}_{min}, \quad (100)$$



^aIf $q_{max} > 0.6\text{\AA}^{-2}$ is imposed, the interval $[q_{min}, q_{max}]$ often includes more than several tens of q -values of diffraction peaks. We should not increase q_{min} , because smaller q -values have better accuracy. As a result, $[q_{min}, q_{max}]$ set by the default parameters often does not satisfy the theoretical requirement $q_{max} > 0.6\text{\AA}^{-2}$. Nevertheless, powder auto-indexing succeeded in all the test data. This is considered to be because all test data satisfy $d \geq 2.7\text{\AA}$ (and many also satisfy $d \geq 4\text{\AA}$), and because the formula (11) of Lagarias *et al.* provides an overestimation of D_3 .

Figure 15: Range $[q_{min}, q_{max}]$ of N_{peak} q -values used in powder auto-indexing.

where 5\AA^3 is chosen as the threshold for the minimum volume of existing crystals, and v_{20} is the upper bound of $\text{Vol}(\mathbb{R}^3/L^*)$, estimated using the 20 smallest elements of $\Lambda^{obs} := \{q_1, q_2, \dots, q_{N_{Peak}}\}$ by:

$$v_j := \frac{2\pi}{3} \frac{q_j^{3/2} - q_1^{3/2}}{j - 1}. \quad (101)$$

Equation (101) is based on the following formula, which holds for any $0 < r < R$.

$$\begin{aligned} & \frac{2\pi(R^3 - r^3)}{3} \left| \{r^2 \leq |l^*|^2 \leq R^2 : l^* \in L^*\} \right|^{-1} \\ & \geq \frac{4\pi(R^3 - r^3)}{3} \left| \{l^* \in L^* : r^2 \leq |l^*|^2 \leq R^2\} \right|^{-1} \rightarrow \text{Vol}(\mathbb{R}^3/L^*) \text{ as } R \rightarrow \infty. \end{aligned} \quad (102)$$

Note that 20 and 30 are chosen empirically, and $v_{N_{Peak}}$ is used instead of v_{20} if $N_{Peak} < 20$. We have found no cases when this $[\text{Vol}_{min}, \text{Vol}_{max}]$ fails to contain the correct volume.

8.2 Results

We prepared 26 + 4 powder diffraction patterns as test data. A summary of the first 26 test data is presented in Table 9. The remaining four are presented in Examples 3–6 to illustrate some rather special cases.

We now evaluate how the method in Section 7.3 improved enumeration times. In Section 7.2, we explained that the time is roughly proportional to N_{zone}^2 . From the formula (97), the computation time of the algorithm in Table 6 is approximately proportional to N_{peak}^4 . Figure 16 illustrates the relation between the following pairs in practical tests:

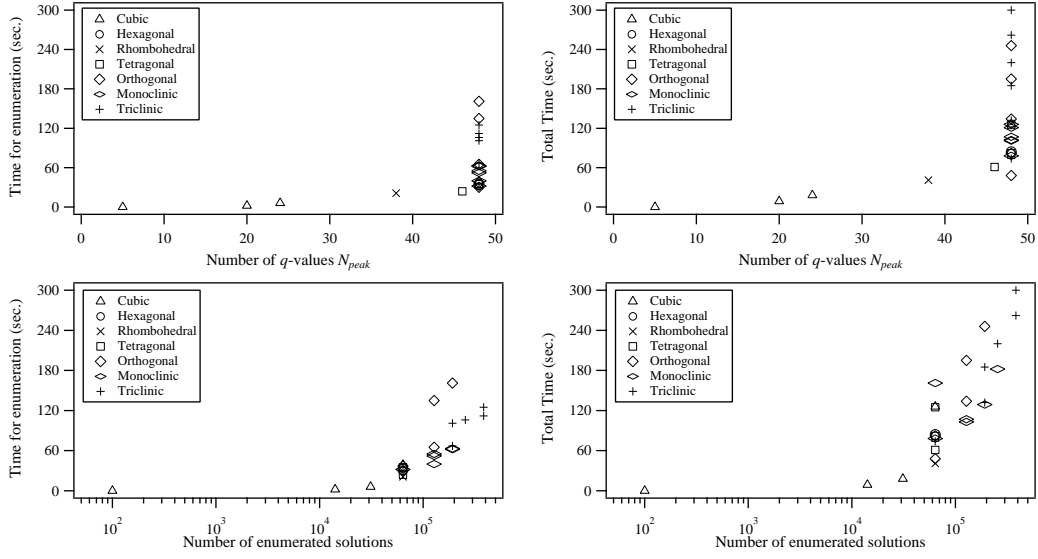
- number of q -values N_{peak} and time for enumeration,
- number of q -values N_{peak} and total time for powder auto-indexing,
- number of enumerated solutions and time for enumeration,

Table 9: Summary of 26 test data^a

Diffractometer	number of patterns
Synchrotron	11
Characteristic X-rays	7
Spallation neutron sources (time-of-flight)	6
Reactor neutron sources	2
Symmetry of lattice	
Triclinic	7
Monoclinic (P)	4
Monoclinic (B)	1
Orthorhombic (P)	5
Tetragonal (P)	2
Tetragonal (I)	1
Rhombohedral	1
Hexagonal	1
Cubic (I)	1
Cubic (F)	3
Distribution of absolute zero-point shifts^b	
– 0.1	17
0.1 – 0.15	1
0.15 –	2

^aAmong eight samples distributed in SDPDRR-2, samples 1–4, 8 are included herein. The result of sample 7 is used in Example 6. Samples 5, 6 are excluded this time, because we could not obtain solutions with de Wolff figures of merit $M_{20} > 10$ using peaks with sufficiently large intensities. These patterns seem to contain a considerable number of false peaks.

^bThe zero-point shifts were computed after execution of powder auto-indexing by non-linear least squares method. Time-of-flight data are excluded here.



^aTotal time includes a) Enumeration of solutions, b) Transformation of every solution to a reduced form, c) Computation of figures of merit, d) Refinement of solutions using linear optimization, e) Error-stable Bravais lattice determination, f) Removal of solutions having a figure of merit less than a (user-input) threshold, g) Removal of duplicate solutions, h) Sorting solutions with a selected figure of merit. Of these, a) and e) are the most time-consuming. The time required for a) consumed approximately half of the total time. We have already contributed to reducing the time for e) in [23].

Figure 16: Computation time of Conograph.

- number of enumerated solutions and total time for powder auto-indexing.

Figure 17 shows the rate of decrease in the size of A_2 as a result of applying the method described in Section 7.3.

Even in the following difficult cases, solutions were obtained without special parameter settings. By Conograph, reliable powder auto-indexing results will become available even for less-experienced users.

Example 3. Non-unique solutions (Figure 18 (a)). For any fixed $C > 0$, the following lattice parameters have exactly the same q -values (cf. 3. in Appendix B).

$$\text{Cubic}(P) : a = b = c = C, \alpha = \beta = \gamma = 90, \quad (103)$$

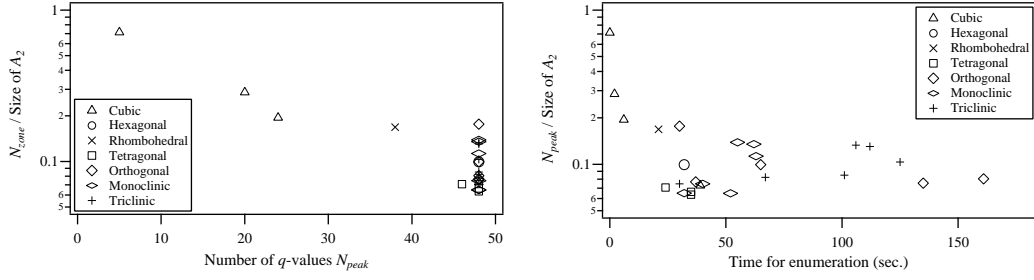
$$\text{Tetragonal}(P) : \sqrt{2}a = \sqrt{2}b = c = C, \alpha = \beta = \gamma = 90. \quad (104)$$

Conograph succeeded in finding both of these.

Example 4. Small q -values are lost (Figure 18 (b)). When the size of the unit cell is large, small q -values are frequently lost because $q < q_{\min}$. We have confirmed that Conograph is very robust to such a loss. This example presents the case in which the 19 smallest non-zero q -values are not included in the observed range $[q_{\min}, q_{\max}]$.

Example 5. Two-phase data (Figure 18 (c)). This powder diffraction pattern is a two-phase sample with the mass ratio 58 : 42. The lattice parameter of the second phase was also enumerated.

Example 6. Poor quality powder diffraction pattern (Figure 18 (d)). This is sample 7 distributed in SDPDRR-2. According to a recent personal report from Le Bail, crystal structures other than sample 7 have been determined. Conograph obtained three reasonable solutions.



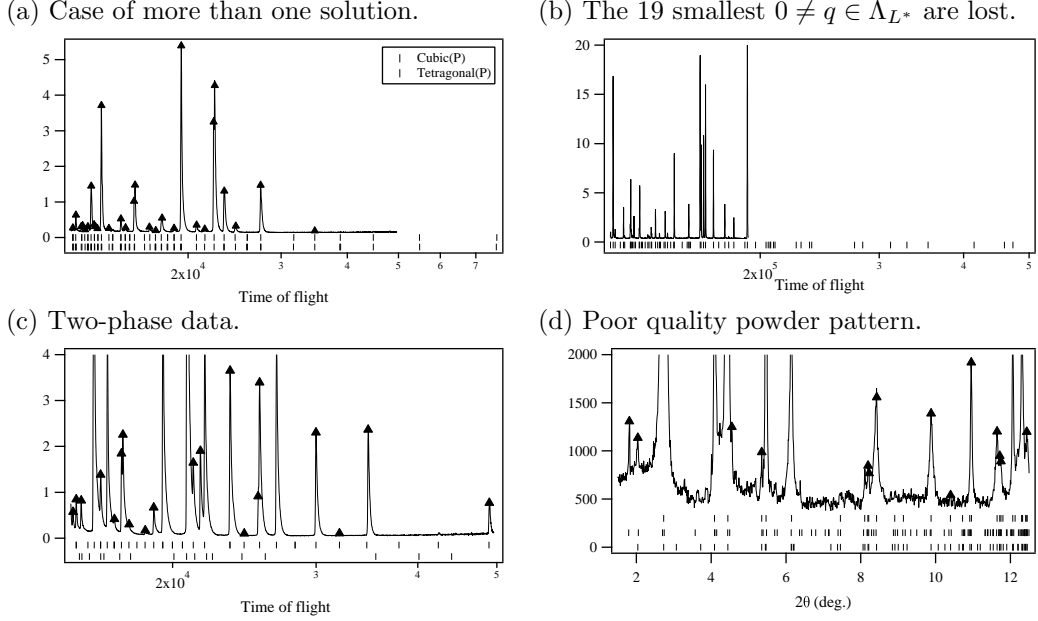
^aExcept for cases of small N_{peak} , all the rates are in the range 0.06–0.18. Under the assumption that the time for enumeration is proportional to N_{zone}^2 , the enumeration is about 32–250 times faster as a consequence. The right-hand figure indicates that a larger improvement was made in more time-consuming cases.

Figure 17: Rates of elements of A_2 used in enumeration algorithm.

9 Conclusion

Powder auto-indexing is divided into two main stages: enumeration and sort of solutions. We contributed mainly to the stage of enumeration by providing a quick and strongly reliable algorithm. For the purpose, Conway’s definition of topographs was generalized to lattices of any rank N using Voronoi’s second reduction theory, holding the association of edges and four lengths $|l_1^*|$, $|l_2^*|$, $|l_1^* + l_2^*|$, $|l_1^* - l_2^*|$ ($l_1^*, l_2^* \in L^*$). By using common properties of systematic absences proved in Theorems 1, 3 and 4, the algorithm is shown to work regardless of the type of systematic absences. These properties are stated as distribution rules for reciprocal lattice vectors corresponding to systematic absences on a topograph. Such rules have not been known so far, and will be also useful in other problems of crystallography. Our enumeration algorithm was implemented in Conograph. Conograph obtained successful results, even for difficult cases. These examples proved that the new enumeration method is robust against missing and false elements in the set of lattice vector lengths extracted from a powder diffraction pattern. Topographs were also utilized to speed up our enumeration algorithm. In practical tests, we found that the improvement reduced the enumeration time to 1/250–1/32.

Acknowledgments The author would like to extend her gratitude to Professor T. Oda of the University of Tokyo for his daily encouragements, to Visible Information Inc. for their cooperation in implementing the Conograph GUI, and to Professor E. Hitzer of the University of Fukui for proposing the impressive name “Conograph.” I would also like to thank Dr. K. Fujii and Professors H. Uekusa, T. Ozeki of the Tokyo Institute of Technology, Dr. S. Torii, Dr. J. Zhang, Dr. M. Ping, and Professors M. Yonemura, T. Kamiyama of KEK, and Professors A. Hoshikawa and T. Ishigaki of Ibaraki University for their valuable comments and for offering test data. This research was partly supported by a Grant-in-Aid for Young Scientists (B) (No. 22740077) and the Ibaraki Prefecture (J-PARC-23D06).



^aTriangles indicate peak positions detected by a peak-search program (corresponding to q -values in Λ^{obs}). Tick marks represent the peak positions corresponding to $q \in \Lambda_{L^*} := \{|l^*|^2 : l^* \in L^*\}$. It is seen that two different lattices have exactly the same q -values. Both parameters are detected by Conograph.

^bThis example presents the case in which the 19 smallest non-zero q -values are not in the observed range. It is confirmed that Conograph is very robust to such a loss. The lattice parameters are $a = 8.82$, $b = 9.76$, $c = 9.78$, $\alpha = \gamma = 90$, $\beta = 104$, and a diffraction pattern from a back-scattering bank of neutron sources is used.

^cThis powder diffraction pattern is a case of a two-phase sample with the mass ratio 58 : 42.

$$\text{Upper} : a = b = c = 12.0 \text{ (Cubic(I), } M_{20} = 7.0), \quad (105)$$

$$\text{Lower} : a = b = 4.8, c = 13.0 \text{ (Rhombohedral, } M_{20} = 1.4). \quad (106)$$

The unit cell of the first phase is much larger than that of the second phase. As a result, the lattice parameter of the first phase gained the best de Wolff figure of merit among all the enumerated solutions, regardless of peaks derived from the second phase. The lattice parameter of the second phase was also enumerated by Conograph. This supports our claim in Section 7.4 that the enumeration process is robust to missing and false elements of Λ^{obs} . However, the figure of merit of the second phase was considerably small, owing to the first-phase peaks. Therefore, it is almost impossible to judge which one is the correct second-phase lattice.

^dThis is sample 7 distributed in SDPDRR-2, held in 2002. The tick marks are peak positions of the following solutions output by Conograph:

$$\text{Upper} : a = 3.99, b = 11.5, c = 17.1, \alpha = 77.9, \beta = 84.6, \gamma = 80.4 \text{ (} M_{20} = 6.6), \quad (107)$$

$$\text{Middle} : a = 4.07, b = 22.5, c = 25.7, \alpha = 88.8, \beta = 87.4, \gamma = 84.8 \text{ (} M_{20} = 7.7), \quad (108)$$

$$\text{Lower} : a = 3.95, b = 17.1, c = 22.8, \alpha = 78.4, \beta = 86.6, \gamma = 84.1 \text{ (} M_{20} = 12.9). \quad (109)$$

The upper lattice parameter is the solution proposed by participants in SDPDRR-2. This regards the two leftmost peaks to be false. The others are newly found by Conograph. The second solution assumes that some peaks are embedded in background noise, and the third is intermediate. This result proves that Conograph makes the stage of enumeration much more reliable. Although the correct solution is provided by a larger M_{20} normally, this is not always right. In particular, it becomes very difficult to select the best solution from peak positions alone, when powder patterns are assumed to have a number of diffraction peaks embedded in background noise or false peaks as intense as diffraction peaks.

Figure 18: Results of Conograph.

A On equivalence between powder diffraction patterns and average theta series

For any periodic function \wp with the period lattice L satisfying $\int_{\mathbb{R}^N/L} |\wp(x)| dx < \infty$, an *average theta series* $\Theta_\wp(z)$ is defined by:

$$\Theta_\wp(z) := \frac{1}{\text{vol}(\mathbb{R}^N/L)} \sum_{l \in L} \int_{(\mathbb{R}^N/L)^2} \wp(x) \wp(y) e^{2\pi\sqrt{-1}z|x-y+l|^2} dx dy, \quad (110)$$

where $\text{vol}(\mathbb{R}^N/L)$ is the volume of \mathbb{R}^N/L . ($\Theta_\wp(z)$ is invariant if a sublattice $L_2 \subsetneq L$ is regarded as the period of the same \wp . This definition is a generalization of the average theta series defined in 2.3, Chapter 2 of [8].) The infinite sum converges uniformly and absolutely in any compact subset of $\{z \in \mathbb{C} : \text{Im}(z) > 0\}$.

Using the Poisson summation formula, the functional equation for $\Theta_\wp(z)$ is obtained:

$$\Theta_\wp(z) = \left(\frac{\sqrt{-1}}{2z} \right)^N \sum_{l^* \in L^*} e^{-\frac{\pi\sqrt{-1}}{2z}|l^*|^2} |\hat{\wp}(l^*)|^2, \quad (111)$$

where $\hat{\wp}(l^*) := \text{vol}(\mathbb{R}^N/L)^{-1} \int_{\mathbb{R}^N/L} \wp(x) e^{-2\pi\sqrt{-1}x \cdot l^*} dx$.

Assuming that $f_{\text{powder}}(q; \wp)$ in (7) equals 0 for any $q < 0$, the Fourier transform of $(2\sqrt{q})^{-1} f_{\text{powder}}(q; \wp)$ is an average theta series.

$$\begin{aligned} \int_{\mathbb{R}} f_{\text{powder}}(q; \wp) e^{2\pi\sqrt{-1}qz} \frac{dq}{2\sqrt{q}} &= \int_{\mathbb{R}} \left(\int_{|x^*|^2=q} \sum_{l^* \in L^*} |\hat{\wp}(l^*)|^2 \delta(x^* - l^*) e^{2\pi\sqrt{-1}qz} dx^* \right) \frac{dq}{2\sqrt{q}} \\ &= \int_{\mathbb{R}^N} \sum_{l^* \in L^*} |\hat{\wp}(l^*)|^2 \delta(x^* - l^*) e^{2\pi\sqrt{-1}|x^*|^2 z} dx^* \\ &= \sum_{l^* \in L^*} e^{2\pi\sqrt{-1}|l^*|^2 z} |\hat{\wp}(l^*)|^2 = (-2\sqrt{-1}z)^{-N} \Theta_\wp \left(-\frac{1}{4z} \right). \end{aligned} \quad (112)$$

Therefore, information obtained from a powder diffraction pattern is theoretically equivalent to that from an average theta series.

B Theorems on the cardinality of solutions

It is well known that the equivalence class of $S_0 \in \mathcal{S}_{>0}^N$ is not uniquely determined, even if all elements of $\Lambda_{S_0} := \{ {}^t v S_0 v : 0 \neq v \in \mathbb{Z}^N \}$ are provided. However, for $N \leq 4$, it is possible to obtain a finite set containing all the equivalence classes of $S \in \mathcal{S}_{>0}^N$ with $\Lambda_S = \Lambda_{S_0}$ (cf. Appendix C).

In this section, several known theorems about the cardinality of solutions are summarized for reference.

- Case $N = 1$.** In this case, Λ_\wp in (8) generates L^* over \mathbb{Z} . (Otherwise, let $L_2^* \subsetneq L^*$ be the lattice generated by Λ_\wp . Then, the reciprocal lattice L_2 of L_2^* is the period lattice of \wp , since $\wp(x) = \sum_{l^* \in L_2^*} \hat{\wp}(l^*) e^{2\pi\sqrt{-1}x \cdot l^*}$ holds. This is a contradiction.) Therefore, the determination of \tilde{L} is straightforward, even from Λ_\wp .
- Case $N = 2$.** It was shown by Delone (and independently by Watson [34], [35]) that, up to a factor, the following is the only pair of inequivalent positive-definite symmetric matrices that have the same representations over \mathbb{Z} .

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}. \quad (113)$$

3. **Case $N = 3$.** From the case of $N = 2$, an infinite family of inequivalent pairs that have the same representations over \mathbb{Z} is obtained.

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & c \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & c \end{pmatrix}. \quad (114)$$

The following is another example (see [20] for the proof).

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \quad (115)$$

A result from Conograph for this case is presented in Example 3 of Section 8.2.

4. **Case $N = 4$.** $S \in \mathcal{S}_{>0}^N$ is called *universal* if Λ_S equals $\mathbb{Z}_{>0}$, the set of all positive integers. It was confirmed by Bhargava and Hanke that the number of equivalence classes of universal $S \in \mathcal{S}_{>0}^4$ equals 6436 [1].
5. **Case $N \geq 5$.** From the existence of universal quadratic forms in $N = 4$, there are infinitely many $S \in \mathcal{S}_{>0}^5$ such that $\Lambda_S = \mathbb{Z}_{>0}$.

C Lattice determination from a complete set of lattice vector lengths

In this section, for any $N \leq 4$ and a given $\Lambda_{S_0} \subset \mathbb{R}_{>0}$ of some $S_0 \in \mathcal{S}_{>0}^N$, an algorithm to enumerate all the equivalence classes of $S \in \mathcal{S}_{>0}^N$ satisfying $\Lambda_{S_0} = \Lambda_S$ is introduced.

First, we recall that $S := (s_{ij})_{1 \leq i, j \leq N} \in \mathcal{S}_{>0}^N$ is *Minkowski-reduced* if and only if S satisfies

$$s_{ii} = \min\{^t v_i S v_i : \{\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, v_i\} \text{ is a primitive set of } \mathbb{Z}^N\}. \quad (116)$$

When S is Minkowski-reduced, the entries of S satisfy

$$0 < s_{11} \leq \dots \leq s_{NN}, \quad 2|s_{ij}| \leq s_{ii} \quad (1 \leq i < j \leq N). \quad (117)$$

Table 10 gives a recursive procedure to generate all the candidates for S from $\Lambda^{obs} := \Lambda_{S_0} \cap (0, c)$, when $0 < c \leq \infty$ is large enough. If the recursive procedure is started with arguments $m = n = I = J = 1$, all positive-definite symmetric matrices $S := (s_{ij})_{1 \leq i, j \leq N}$ satisfying the followings are enumerated in an array Ans .

$$\begin{cases} s_{11} \leq \dots \leq s_{NN}, \quad 2|s_{ij}| \leq s_{ii} \quad (1 \leq i < j \leq N), \\ s_{ii+1} \leq 0 \quad (1 \leq i < N), \\ \{s_{ii} : 1 \leq i \leq N\} \cup \{s_{ii} + s_{jj} + 2s_{ij} : 1 \leq i < j \leq N\} \subset \Lambda^{obs}. \end{cases} \quad (118)$$

Consequently, any Minkowski-reduced S satisfying (118) and $3s_{NN} \leq q_M$ is output in Ans . As a result of Proposition C.1, the recursive procedure is always completed in a finite number of steps, even if we set $c = \infty$, i.e., if Λ_{S_0} is used instead of Λ^{obs} . This indicates all the equivalence classes of $S \in \mathcal{S}^N$ satisfying $\Lambda_S = \Lambda_{S_0}$ are enumerated by the recursive procedure, if sufficiently large c is selected. Consequently, the number of equivalence classes of $S \in \mathcal{S}_{>0}^N$ satisfying $\Lambda_S = \Lambda_{S_0}$ is finite for any $\Lambda_{S_0} \subset \mathbb{R}_{>0}$.

In the remainder of this section, we give a proof of Proposition C.1.

Proposition C.1. *Suppose that $S \in \mathcal{S}_{>0}^N$ and $S_2 \in \mathcal{S}_{>0}^{N_2}$ and $0 < N_2 < N \leq 4$. Then, $\Lambda_{S_2} \not\subset \Lambda_S$.*

Table 10: Recursive procedure for lattice determination from a perfect set of representations.

		void enumerateLattice	$(N, \Lambda^{obs}, m, n, S, I, J, Ans)$
(Input)	$1 \leq N \leq 4$:	number of rows and columns of $S_0 \in \mathcal{S}_{>0}^N$,
	$\Lambda^{obs} := \langle q_1, \dots, q_M \rangle \neq \emptyset$:	a sorted sequence of all the elements of Λ_{S_0} that belong to the interval $(0, c)$,
	m, n	:	integers $1 \leq m \leq n \leq N$ indicating the (m, n) -entry of S ,
	$S := (s_{ij})$:	$N \times N$ positive-definite symmetric matrix that fulfills $0 < s_{11} \leq \dots \leq s_{nn}$, $2 s_{ij} \leq s_{ii}$ ($1 \leq i < j < n$), $s_{ii+1} \leq 0$ ($1 \leq i \leq n$).
	I, J	:	integers indicating
			$\begin{cases} s_{nn} \in \{q_i : I \leq i \leq J\} & \text{if } m = n, \\ s_{nn} = q_I \text{ and } s_{mm} + s_{nn} + 2s_{mn} \in \{q_i : I \leq i \leq J\} & \text{otherwise.} \end{cases}$
(Output)	Ans	:	array of positive-definite $N \times N$ symmetric matrices.
1:	(start)	for $l = I$ to J do	
2:		if $m = n$ then	
3:		$s_{nn} = q_l$.	
4:		else	
5:		$s_{mn} = s_{nm} = \frac{1}{2}(q_l - s_{mm} - s_{nn})$.	
6:		end if	
7:		if $m = 1$ then	
8:		if $\det(s_{ij})_{1 \leq i, j \leq n} > 0$ then	
9:		if $n \geq N$ then	
10:		Insert S in Ans .	
11:		else	
12:		$m_2^a := \max \left\{ 1 \leq i \leq M : \begin{array}{l} q_1, \dots, q_{i-1} \text{ are representations of} \\ \text{the submatrix } (s_{ij})_{1 \leq i, j \leq n} \text{ over } \mathbb{Z} \end{array} \right\}$.	
13:		if $m = n$ then	
14:		Call searchLattice($N, \Lambda^{obs}, n + 1, n + 1, S, l, m_2, Ans$).	
15:		else	
16:		Call searchLattice($N, \Lambda^{obs}, n + 1, n + 1, S, I, m_2, Ans$).	
17:		end if	
18:		end if	
19:		end if	
20:		else	
21:		if $m = n$ then	
22:		$m_2 := \max \{ 1 \leq i \leq M : q_i \leq s_{n-1, n-1} + s_{nn} \}$.	
23:		Call searchLattice($N, \Lambda^{obs}, m - 1, n, S, l, m_2, Ans$).	
24:		else	
25:		$m_2 := \max \{ 1 \leq i \leq M : q_i \leq 2s_{m-1, m-1} + s_{nn} \}$.	
26:		Call searchLattice($N, \Lambda^{obs}, m - 1, n, S, I, m_2, Ans$).	
27:		end if	
28:		end if	
29:		end for	

^aProposition C.1 claims that there always exists $m_2 < \infty$, even if the sequence Λ^{obs} is replaced by Λ_{S_0} virtually.

Lemma C.1 is utilized in the proof.

Lemma C.1. Any $S \in \mathcal{S}_{>0}^N$ is represented as a finite sum $\sum_k \lambda_k S_k$ such that every S_k is a positive-definite symmetric matrix with rational entries, and $\lambda_k \in \mathbb{R}_{>0}$ are linearly independent over \mathbb{Q} .

Proof. For any $S := (s_{ij})_{1 \leq i, j \leq N} \in \mathcal{S}^N$, let v_S be the vector ${}^t(s_{11}, s_{12}, \dots, s_{ij}, \dots, s_{NN})$ of length $\frac{N(N+1)}{2}$. Then, \mathcal{S}^N is identified in $\frac{N(N+1)}{2}$ -dimensional vector space by the map $S \mapsto v_S$. Define a set P by:

$$P := \{I \subset \{s_{ij} : 1 \leq i \leq j \leq N\} : s_{ij} \in I \text{ are linearly independent over } \mathbb{Q}\} \quad (119)$$

Then, P is not empty. Let $\{t_1, \dots, t_m\}$ be one of the maximal elements of P under inclusive order. When vectors ${}^t(t_1, \dots, t_m)$ and ${}^t(1, \dots, 1)$ of length m are denoted by \mathbf{t} and $\mathbf{1}_m$ respectively, there exists an $\frac{N(N+1)}{2} \times m$ rational matrix C such that $v_S = C\mathbf{t}$. Furthermore, there exists $\epsilon > 0$ such that for any $U := (u_{ij}) \in GL_m(\mathbb{R})$ with entries $|u_{ij}| < \epsilon$, every column of $C(\mathbf{t}\mathbf{1}_m - U)$ is the image of a positive-definite symmetric matrix by the map $S \mapsto v_S$.

Let $\mathbf{1}_m$ be the identity matrix of size m . If ${}^t\mathbf{1}_m U^{-1} \mathbf{t} \neq 1$, we have equations:

$$(\mathbf{t}\mathbf{1}_m - U)^{-1} = U^{-1}(({}^t\mathbf{1}_m U^{-1} \mathbf{t} - 1)^{-1} \mathbf{t}\mathbf{1}_m U^{-1} - \mathbf{1}_m), \quad (120)$$

$$(\mathbf{t}\mathbf{1}_m - U)^{-1} \mathbf{t} = ({}^t\mathbf{1}_m U^{-1} \mathbf{t} - 1)^{-1} U^{-1} \mathbf{t}. \quad (121)$$

If all the entries of $U^{-1} \mathbf{t}$ are negative, we have that ${}^t\mathbf{1}_m U^{-1} \mathbf{t} < 0$, and every entry of $(\mathbf{t}\mathbf{1}_m - U)^{-1} \mathbf{t}$ is positive. Clearly, there exists $U := (u_{ij}) \in GL_m(\mathbb{R})$ such that $|u_{ij}| < \epsilon$, every entry of $U^{-1} \mathbf{t}$ is negative, and the matrix $\mathbf{t}\mathbf{1}_m - U$ is rational. Fix such a U .

Let S_k ($1 \leq k \leq m$) be a positive-definite symmetric matrix satisfying $C(\mathbf{t}\mathbf{1}_m - U) = (v_{S_1}, \dots, v_{S_m})$. S is then represented as a linear sum of rational S_k with positive coefficients as follows:

$$v_S = C\mathbf{t} = (v_{S_1}, \dots, v_{S_m})(\mathbf{t}\mathbf{1}_m - U)^{-1} \mathbf{t}. \quad (122)$$

Hence, the statement is proved. \square

For any ring $R_2 \subseteq R$ and a symmetric matrix S with entries in R_2 , let $\Lambda_S(R)$ be the set $\{{}^t v S v : 0 \neq v \in R^N\}$ consisting of representations of S over R . If $0 \in \Lambda_S(R)$, S is said to be *isotropic* over R . Otherwise, S is *anisotropic* over R .

Proof of Proposition C.1. The statement holds if it is true when $N_2 + 1 = N = 4$. By Lemma C.1, it is sufficient if the statement is proved in the case that S, S_2 are rational. Any non-singular quadratic form over \mathbb{Q}_p of rank 4 satisfies $\Lambda_S(\mathbb{Q}_p) \supset \mathbb{Q}_p^\times$ for any p . On the other hand, any anisotropic quadratic form over \mathbb{Q}_p of rank 3, there exists a finite prime p such that $\Lambda_{S_2}(\mathbb{Q}_p) \not\supset \mathbb{Q}_p^\times$, (cf. Corollary 2 of Theorem 4.1 in Chapter 6, [5]). If $\Lambda_{S_2} \supset \Lambda_S$, $\Lambda_{S_2}(\mathbb{Q}) \supset \Lambda_S(\mathbb{Q})$, therefore $\Lambda_{S_2}(\mathbb{Q}_p) \supset \Lambda_S(\mathbb{Q}_p)$ is required for any p . This is a contradiction. \square

References

- [1] M. Bhargava and J. Hanke. Universal quadratic forms and the 290 theorem. To appear in Invent. Math.
- [2] L. Bieberbach. Über die bewegungsgruppen der euklidischen räume. *Mathematische Annalen*, 70(3):297–336, 1911.

- [3] L. Bieberbach. Über die bewegungsgruppen der euklidischen räume (zweite abhandlung.) die gruppen mit einem endlichen fundamentalbereich. *Mathematische Annalen*, 72(3):400–412, 1912.
- [4] A. Boultif and D. Louër. Powder pattern indexing with the dichotomy method. *J. Appl. Cryst.*, 37:724–731, 2004.
- [5] J. W. S. Cassels. *Rational Quadratic Forms*. Academic Press, London/New York, 1978.
- [6] J. H. Conway. *The sensual (quadratic) form*. Carus Mathematical Monographs 26, Mathematical Association of America, 1997.
- [7] J. H. Conway and N. J. A. Sloane. Low-dimensional lattices vi: Voronoi reduction of three-dimensional lattices. *Proc. Royal Soc. London, Series A*, 436:55–68, 1992.
- [8] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups (3rd ed.)*, volume 290 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1998.
- [9] P. M. de Wolff. On the determination of unit-cell dimensions from powder diffraction patterns. *Acta Cryst.*, 10:590–595, 1957.
- [10] P. M. de Wolff. Detection of simultaneous zone relations among powder diffraction lines. *Acta Cryst.*, 11:664–665, 1958.
- [11] P. M. de Wolff. A simplified criterion for the reliability of a powder pattern indexing. *J. Appl. Cryst.*, 1:108–113, 1968.
- [12] P. Engel, T. Matsumoto, G. Steinmann, and H. Wondratschek. The non-characteristic orbits of the space groups. *Z. Kristallogr., Supplement Issue No. 1*, 1984.
- [13] P. M. Gruber and S. S. Ryshkov. Facet-to-facet implies face-to-face. *European Journal of Combinatorics*, 10:83–84, 1989.
- [14] T. Hahn. *International tables for Crystallography*, volume A. Dordrecht:Kluwer, 1983.
- [15] T. Ito. A general powder x-ray photography. *Nature*, 164:755–756, 1949.
- [16] M. Kac. Can one hear the shape of a drum? *American Mathematical Monthly*, 73(4):1–23, 1966.
- [17] F. Kohlbeck and E.M. Horl. Trial and error indexing program for powder patterns of monoclinic substances. *J. Appl. Cryst.*, 11:60–61, 1978.
- [18] J. C. Lagarias, H. W. Lenstra, JR., and C. P. Schnorr. Korkin-zolotarev bases and successive minima of a lattice and its reciprocal lattice. *COMBINATORICA*, 10(4):333–348, 1990.
- [19] A. Le Bail. Monte carlo indexing with McMaille. *Powder Diffraction*, 19:249–254, 2004.
- [20] Y. S. Moon. Universal quadratic forms and the 15-theorem and 290-theorem (thesis). 2008.
- [21] M. A. Neumann. X-cell: a novel indexing algorithm for routine tasks and difficult cases. *J. Appl. Cryst.*, 36(4):356–365, 2003.

- [22] P. Niggli. *Kristallographische und strukturtheoretische Grundbegriffe. Handbuch der Experimentalphysik*, volume 7. Leipzig: Akademische Verlagsgesellschaft, 1928.
- [23] R. Oishi-Tomiyasu. Rapid bravais-lattice determination algorithm for lattice parameters containing large observation errors. *Acta Cryst. A.*, 68:525–535, 2012.
- [24] S. S. RYSHKOV and E. P. BARANOVSKII. C-types of n-dimensional lattices and 5-dimensional primitive parallelohedra (with application to the theory of coverings). *Proceedings of the Steklov Institute of Mathematics*, 137, 1976.
- [25] A. Schiemann. Ein beispiel positiv definiter quadratischer formen der dimension 4 mit gleichen darstellungszahlen. *Archiv der Mathematik*, 54:372–375, 1990.
- [26] A. Schiemann. Ternary positive definite quadratic forms are determined by their theta series. *Mathematische Annalen*, 308:507–517, 1997.
- [27] E. Selling. Über die binären und ternären quadratischen formen. *J. Reine Angew. Math.*, 77:143–229, 1874.
- [28] J. P. Serre. *Corps Locaux*. Hermann, Paris, 1968.
- [29] D. H. Templeton. Systematic absences corresponding to false symmetry. *Acta Cryst.*, 9:199–200, 1956.
- [30] F. Vallentin. *Sphere coverings, lattices, and tilings (in low dimensions)*. PhD thesis, Technical University Munich, Germany, 2003.
- [31] J. W. Visser. A fully automatic program for finding the unit cell from powder data. *J. Appl. Crystallogr.*, 2:89–95, 1969.
- [32] G. F. Voronoi. Sur quelques proprietes des formes quadratiques positives parfaites. *J. Reine Angew. Math.*, 133:97–178, 1907.
- [33] G. F. Voronoi. Nouvelles applications des parametres continus a la theorie des formes quadratiques. *J. Reine Angew. Math.*, 134:198–287, 1908.
- [34] G. L. Watson. Determination of binary quadratic form by its values at integer points. *MATHEMATIKA*, 26:72–75, 1979.
- [35] G. L. Watson. Determination of binary quadratic form by its values at integer points: Acknowledgement. *MATHEMATIKA*, 27:188, 1980.